TECHNOMETRICS

A Journal of Statistics for the Physical, Chemical and Engineering Sciences

Engin. Library
QA
276
.A1
T25
v. 3
no. 1

FEBRUARY 1961

VOLUME 3 NUMBER 1

Published quarterly by:

The American Society for Quality Control and The American Statistical Association

THE UNIVERSITY OF MICHIGAN LIBRARIES

TECHNOMETRICS

TECHNOMETRICS

The purpose of **Technometrics** is to contribute to the development and use of statistical methods in the physical, chemical and engineering sciences. This objective places a high premium on succinct communication among the physical scientist, engineer, statistician and mathematician. The journal will accept for publication papers describing new statistical techniques expected to be useful in these sciences, papers illustrating the application of known statistical methods to new or novel environments, expository or tutorial papers on particular statistical methods, and papers dealing with the philosophy and problems of applying statistical methods to research, development, design and performance. Brief descriptions of problems requiring solution and short technical notes will also be accepted for publication. Letters to the Editor, signed by the author and limited in length will be published when they are considered timely and appropriate. All papers should contain a short but clear summary of contents and conclusions, an expository section containing numerical examples whenever practicable, and appropriate additional sections relating to technical derivations.

Subscription Rates

The annual subscription rate for members of the American Society for Quality Control or the American Statistical Association is $6.00 a year. The annual subscription rate for non-members is $8.00 a year.

Members of the sponsoring societies, the American Society for Quality Control and the American Statistical Association, may subscribe to the journal while paying their annual dues or by check or money order made out to Technometrics and mailed to:

TECHNOMETRICS
Post Office Box 587
Benjamin Franklin Station
Washington 4, D. C.

All non-member subscriptions should be mailed to this address. Communications concerning changes of address, subscriptions, back numbers, etc., should also be sent to this address. Whenever possible, a copy of the address taken from an issue of the periodical should accompany a change of address request. All subscription fees are payable in the currency of the United States of America.

Communication concerning *membership* in either of the sponsoring societies should be sent to that society.

Second class postage paid at Richmond, Va.

TECHNOMETRICS

A Journal of Statistics for the Physical, Chemical and Engineering Sciences

Published Quarterly by

THE AMERICAN SOCIETY FOR QUALITY CONTROL

and the

AMERICAN STATISTICAL ASSOCIATION

Editor

J. STUART HUNTER

Associate Editors

WILLIAM ALLEN
G. A. BARNARD
C. A. BENNETT
CUTHBERT DANIEL
BESSE B. DAY
R. J. HADER
DAVID R. HOWES
FRED C. LEONE
MARTIN WILK
MARVIN ZELEN

Management Committee

PAUL S. OLMSTEAD, *Chairman*

For the American Society for Quality Control

WARREN R. PURCELL
MAYNARD RENNER
H. L. WEHRLY

For the American Statistical Association

IRVING BURR
ALMARIN PHILLIPS
DONALD C. RILEY

Published Quarterly in February, May, August and November by the Technometrics Management Committee of the American Society for Quality Control and the American Statistical Association. Editorial Office: Mathematics Research Center, U. S. Army; The University of Wisconsin, Madison, Wisconsin. Publication Office: Wm. Byrd Press, P. O. Box 2W, Richmond 5, Virginia. Second class mailing privileges granted at Richmond, Virginia.

Composed and Printed at the

WILLIAM BYRD PRESS, INC., RICHMOND, VIRGINIA, U.S.A.

CONTENTS

TECHNOMETRICS, Vol. 3, No. 1, FEBRUARY, 1961

Cumulative Sum Charts ..... *E. S. Page* 1

Average Run Lengths in Cumulative Chart Quality Control Schemes ..... *P. L. Goldsmith and H. Whitfield* 11

Prediction Regions for Several Predictions from a Single Regression Line ..... *Gerald J. Lieberman* 21

The Robustness of Life Testing Procedures Derived from the Exponential Distribution *Marvin Zelen and Mary C. Dannemiller* 29

An Application of a Balanced Incomplete Block Design *Peter W. M. John* 51

Multi-component Systems and Structures and Their Reliability. . *Z. W. Birnbaum, J. D. Esary and S. C. Saunders* 55

An Asymptotic Distribution for an Occupancy Problem with Statistical Applications . . . . *M. Halperin and G. L. Burrows* 79

Outliers in Patterned Experiments: A Strategic Appraisal *Irwin D. J. Bross* 91

Multiple Comparisons with a Control in Balanced Incomplete Block Designs ..... *D. S. Robson* 103

Estimates of Bounded Relative Error for the Mean Life of an Exponential Distribution ..... *Benjamin Epstein* 107

On Methods of Constructing Sets of Mutually Orthogonal Latin Squares Using a Computer. II *R. C. Bose, I. M. Chakravarti and D. E. Knuth* 111

Book Reviews ..... *Acheson J. Duncan and William Beranek* 119

Statistical Programs for High Speed Computers 123

Notices 127

VOL. 3, NO. 1 TECHNOMETRICS FEBRUARY, 1961

Cumulative Sum Charts

E. S. PAGE

University of Durham Computing Laboratory, Newcastle upon Tyne, England

This paper, presented orally to the Gordon Research Conference on Statistics in Chemistry in July 1960, traces the development of process inspection schemes from the original methods of Shewhart to the new charts using cumulative sums, and surveys the present practice in the operation of schemes based on cumulative sums. In spite of the completely different appearance of the visual records kept for Shewhart and cumulative sum charts, a continuous sequence of development from the one type of scheme to the other can be established. The criteria by which a particular process inspection scheme is chosen are also developed and the practical choice of schemes is described.

INTRODUCTION

The type of chart that I want to discuss is one which is to be used for process inspection and on which is plotted the cumulative sum of the observations themselves or of some function of the observations. Barnard (2) has called this plotting of cumulative sums "a fundamental change in the classical procedure" and so perhaps it is—it certainly seemed so in 1954 (9). In spite of this, however, it is now possible to trace its close connection with the original Shewhart charts through the various modifications that have been made. I should like to go over this development before talking about cumulative sum charts themselves and their possible future developments.

Process inspection means different things to different people; I interpret the principal aim of process inspection to be the furnishing of information to us, either to assure us that the process is producing its output in the specified manner, or to warn us promptly that some departure from specifications is occurring so that some action may be taken. This action may take many forms; it may be remedial, and perhaps also determine the destination of the recent output; or it might trigger an investigation for any assignable causes of deterioration in quality. In many applications the information from the inspection will be required to indicate the type of action that is necessary, or the amount of it. In particular, the information from the inspection may be required to give an estimate of the amount of departure from specification that is occurring. If the process is always adjusted by an amount dependent on such an estimate the scheme is acting as a servo-mechanism governing the control of the process and it will be called one of process control and the term process inspection will be reserved for the earlier use.

In addition to this principal aim, schemes of process inspection have important applications when the history of the process is to be examined. It is hoped and expected that the method of display of the information obtained will enable

the occurrences of trouble in the process to be readily identified in time and so to make easier the isolation of the causes of trouble. At the other end of the scale too, the display should similarly draw attention easily to any marked improvements in the process and so to isolate the causes of these. We are quite as interested in knowing how to avoid deteriorations in quality as we are in learning how to maintain any improvements that are noticed. While I regard the first aim as the main one, the second is so important that any method of presentation of the information which did not allow the latter would seem to be hardly acceptable.

Shewhart Charts

The type of control chart for process inspection which has been used over some thirty years is basically that suggested by Shewhart (15) with various modifications and amendments that have been made since the original proposal. For these charts, results of individual samples are plotted on a diagram and rules are given to enable an interpretation to be made easily. Possibly the principal value in the introduction of the charts was to emphasize the value of a good visual presentation reinforced by some sensible rules-of-thumb for making decisions based upon the data. In the original proposals the charts were furnished with what we may call action lines, so that when any point fell outside these lines, some action on the part of those tending the process was called for. Clearly, with such a simple rule, the samples themselves were being classed as either good or bad; the good ones where the points fell within the action lines would suggest that the process should continue without special attention or interruption, while the bad ones with points outside the lines demanded the corrective action. The main point to be decided for the operation of such a process inspection scheme was the position at which the action lines should be drawn once the sample size had been determined; the positions that were most frequently adopted were the so called three-sigma limits. For example, on a chart for controlling the mean dimension of a process at a target value μ, the lines will be drawn at $\mu \pm 3.09\ \sigma/\sqrt{n}$ for samples of size n. The lines were to be drawn in these positions because if the process mean were unchanged, the probability that the sample mean of n observations would fall outside the lines and cause action to be taken is 0.002, so that action would be taken unnecessarily on the average only once in every five hundred times. We can see easily now how related the choice of this position for the action lines was to current practice in the testing of hypotheses and whatever comments and criticisms that we may make of this choice now, the pragmatic justification has been enormous. The contribution of Shewhart's charts to the development and satisfactory control of mass production methods must have been vital, irrespective of the monetary savings which these charts have brought to those using them in many other different applications. When we have to decide on the positions of action lines now, we would wish not to rely on a three-sigma dogma but instead to think about the costs and benefits of drawing the action lines at different positions, (10). We should not immediately import the positions more appropriate to other sets of circumstances. As process inspection is a continuous proceedings

being carried out as long as the process is operated, we shall wish to consider the effect of the size of samples to be taken and the frequency of taking them upon the disadvantages and costs of failing to notice immediately a departure from the specification manufacturing conditions, and the disadvantages on the other hand of witch-hunting in a process that is really operating satisfactorily, but has by chance produced a bad sample.

In 1950, Aroian and Levene (1) considered afresh the sort of measures that one should use for assessing the statistical properties of process inspection schemes. If there is an abrupt change in the quality of the product or in the departure from specification we need to know the distribution of the amount produced by the process before the deterioration is noticed by the inspection rule we are using. From this distribution a detailed study of the costs involved could be made in any particular case, but for a general comparison of inspection schemes, something rather simpler is preferable. Such a function is the average amount produced before action is demanded by the inspection scheme and if the rate at which the process is sampled is constant, this is proportional to the average number of items sampled before the action is demanded. The name suggested to me by Anscombe for the latter quantity was the Average Run Length (A.R.L.) in preference to Aroian's slightly longer name. Barnard and Ewan and Kemp take the A.R.L. to be the average number of samples before action is demanded. If items are sampled one by one the two uses are identical, but if each sample contains more than one item they differ. As the available tables for the normal distribution follow Barnard's use we shall adopt it here and note that the tables for fraction defective (9) follow the earlier definition. If the model for the change in process conditions is more complex, it may be essential to do a detailed cost analysis as, for example, Duncan (4) has done in some cases.

The motivation for the modifications that have been suggested to the original Shewhart scheme can be interpreted with reference to such a measure of an inspection scheme's performance as the A.R.L. The inspection scheme should be able to call attention quickly to a serious deterioration in product quality and also to notice before too long any slight but sustained deterioration. A slight deterioration would be shown on the Shewhart chart with action lines by sequences of points departing consistently from the target value, but perhaps insufficiently extreme to fall outside the action lines. Consequently, warning lines were proposed which would be drawn within the action lines at perhaps the two sigma limits,

$$\text{e.g. for the mean } \mu \pm 1.96\sigma/\sqrt{n}$$

with an addition to the rule that if K out of the last N points fall between the warning and action lines, then an investigation would be demanded. This sort of scheme has a steeper average run length function when plotted against the size of the departure from target value than the original Shewhart scheme, (12). Some other rules use runs of points on the control chart, for example, Moore (8) Weiler (19–21) consider rules of the type, "If K consecutive points on the chart fall outside control lines, take action".

Schemes like this have certain advantages over those based on one sample

point only. They have the practical one of requiring only small samples; and they detect large changes with the speed of small samples and small changes as well as the single point rule with large samples. There has been a certain amount of work (10, 19, 20) on the best choice of rule and the positions of the control lines.*

It seems to me that these modifications have been prompted by the realization that the charts contained much more information than was being used by a one-point rule. Instead these new rules were designed to take into account part of the information contained in the chart for a fixed number of samples in the immediate past.

The next step—which does not seem such a big one since Wald's introduction of sequential methods in hypothesis testing—is to relax the restriction that only a fixed number of past samples should be considered in taking the decision. At the same time, if it can be done easily we would like a scheme which will take into account the actual positions of the points on the chart and not merely the classifications into which division of the chart the point falls. The cumulative sum charts do indeed possess these properties and to show them in the simplest possible way I want to consider controlling the mean of a distribution against one-sided deviations from the target value only. Practical cases might be the proportion of defective items or the range of a sample of given size—I know that we shall often want to know about deviations in the "improvement" direction but for now I consider only the departure in one direction—a deterioration.

Suppose that we wish to control the population mean at a target value μ and that we take observations x_i in the course of our process inspection. We then plot

$$S_r = \sum_{i=1}^{r} (x_i - k)$$

on the chart where k is a "reference value" related to the target value μ. It will be such that the mean path is downwards (or horizontal) for good quality and upwards for bad. We will take $k = \mu$ here and return to its choice later. Naturally we only need to add the deviation from the reference value to the last point recorded each time. With this chart we have immediately a picture of the behaviour of the sample means stretching back over the last point, last 2, 3, 4—points as far as we wish. All we need to do is to look at the slopes of the lines joining the current point to the sample points as far back as we like. If we are concerned about departures from specifications greater than the target value, we want to know when the cumulative sum path on the chart takes an upward turn. The easiest type of action rule to adopt is "Take action if the current sample point rises more than a stated amount, h, above the previous lowest point of the sample path." Thus this rule takes into account the actual sizes of the departures and also makes immediately apparent all the recent evidence about a possible change in the population mean.

---

* Work similar to that in (12) is reported by S. W. Roberts, *Bell System Technical Journal*, *37*, 1958, 83.

Before leaving the Shewhart type of chart there is another connection between these charts and cumulative sum charts that I want to mention. When there are both warning and action lines on a Shewhart chart the samples are really being classified as "Bad" ≡ point beyond the action lines, "Doubtful" ≡ point between action and warning lines, "Satisfactory" ≡ within the warning lines. Under the rule "Take action if one point falls outside the action lines or K out of the last N points fall between control lines and action lines" we are demanding action if one "Bad" sample occurs *or* if too many "Doubtful" ones occur within the last N samples. For the application of these rules it is only necessary to count the number of points in the different regions, but in order to see which of these rules is equivalent to the cumulative sum procedure, it is convenient to introduce a scoring system. As an example, consider the rule where $K = N = 3$ so that three consecutive ("doubtful") points outside the warning lines give the same signal as one outside the action lines. If we assign scores to the samples $x_i = -2, +1, +3$ for "Good," "Doubtful" and "Bad" points respectively the rule "Take action if the cumulative sum $\sum_{i=1}^{r} x_i$ rises 3 or more above its minimum" is equivalent in every way to the rule expressed in terms of warning lines. It is easily seen that occurrence of a "Good" point at least pulls the cumulative sum path to its previous minimum and so satisfies the conditions of the rule about consecutive doubtful points only being regarded as important. More generally it can be shown (12) that warning line rules of the above type for $K = 2$, any N (scoring system e.g. "Good" $\equiv -1$, "Doubtful" $\equiv N - 1$, "Bad" $\equiv N$) and for $K = N$ ("Good" $\equiv -(N-1)$, "Doubtful" $\equiv 1$, "Bad" $\equiv N$) are equivalent to cumulative sum rules applied to the scores. It follows that many of the amended Shewhart rules that are now in use are in fact equivalent to a restricted type of cumulative sum rule. However, the full advantages are not being obtained; the use of the actual observations and the easy and vivid visual picture of a cumulative sum chart are both missing.

Cumulative Sum Charts

The operation of cumulative sum charts in practice is very similar to the operation of the usual charts. The differences are in the type of visual record made and the criteria for deciding to take action. It is necessary to take some care with the plotting of the cumulative sums because they can be extravagant with the paper—paper is cheap but if the record covers too great an area some inconvenience results. This can be easily avoided with charts for one-sided deviations; suppose we wish to detect a positive shift in the process parameter—the mean say. Then the mean path of the sum

$$S_r = \sum_{i}^{r} (x_i - k)$$

will take a turn upwards if the process mean increases above k. Any negative increments, $x_i - k < 0$, and parts of the path pointing downwards will give little indication of an increase. As long as the path is tending downwards we are satisfied with the value of the process mean and there is no need to keep the cumulative sum going more and more negative. If we start at $S_0 = 0$, we

can record S_r as long as it is positive and plot it on the chart, taking action if the Alarm value, h is reached; when S_r falls below zero, start the cumulative sum afresh at zero, and again restart at zero after action has been taken. This method of plotting the chart is using the identity of the rule with the repeated application of a Wald sequential test with boundaries at 0, h and with starting point always on the acceptance boundary. We keep on applying the test until one ends at the upper boundary and then take action. (Incidentally, this points out another connection with the Original Shewhart Chart—that was equivalent to the repeated application of fixed sample size tests).

The other quantity needed is to decide what rise, h, in the cumulative sum path should demand action. If the statistical model of the production process is known both in its normal target state and all its possible departure states it would be possible to derive mathematically the optimum value of h. These conditions of full knowledge are likely to be met only rarely and instead we need to look for values of h likely to be satisfactory if not best under quite wide ranges of circumstances. Of course, one value of h will not necessarily suffice for all applications, but we can take heart from the experience with the classical control charts that the three-sigma limits which Shewhart had the wisdom to suggest served so well for so many people. If we want to proceed in this way we can select our schemes based on their average run length function. For example, when the process is in control and samples are independent, a Shewhart chart with just three-sigma limits has average run length 500; this corresponds to an unnecessary interruption of a controlled process on the average every 500 samples.

Charts for two-sided deviations need to take into account departures of the cumulative sum from the mean path in both upward and downward directions. Originally I suggested that we should apply two one-sided schemes simultaneously, one to detect an increase and the other a decrease, but the difficulty of finding the A.R.L. caused me to look for alternative procedures. However, Barnard has re-presented the scheme using a masking device and Goldsmith and Whitfield have evaluated by a Monte Carlo method the A.R.L.'s for several schemes (2), (6). The method is as follows:

Plot the cumulative sum $S_r = \sum_1^r (x_i - \mu)$ where μ is the target value. Then place on the chart a V-shaped mask with an angle 2θ, its vertex a distance d from the current point and its axis horizontal. If all the previous sum path is visible, continue without interrupting the process. If part of the path disappears under the upper arm of the mask then the process mean has *decreased* below its target value. The two quantities giving the shape and position of the mask, θ and d can be chosen to give the scheme desired characteristics. We return to the method of choice in the next section.

These cumulative sum schemes are much more sensitive than the ordinary Shewhart ones, especially to moderate deviations from the target. For the same A.R.L. at the target value the A.R.L. of a cumulative sum scheme at a deviation of about one S.D. can be one quarter of that of Shewhart scheme. This sensitivity is an advantage for processes needing precise control, but is a definite drawback where some slack in the process is permissible. In these cases, cumulative sum charts should *not* be used or should be used with greater

thought. It is easy to see that continued production at a value away from the nominal target will cause much more frequent interruption of the process by the sum chart although the quality of production is acceptable; to some extent the nuisance can be avoided by choosing a scheme with a very large A.R.L. at a nominal target value in the centre of the "slack" region.

Choice of Scheme

The use of three-sigma limits with Shewhart charts became a dogma and it would be unfortunate if a similar sort of dogma were to arise for cumulative sum charts. As the employees of one world-known company are urged, the best way is to "Think". Following this injunction, we need first to decide upon two values of the process parameter μ_0 , μ_1 ; μ_0 is the target value and is that value at which we regard the process as operating satisfactorily and we consequently wish only rare interference with the process; and μ_1 , a value of the process parameter that is regarded as bad and should be noticed quickly. Next we should decide upon the desired values of the A.R.L. L_0 and L_1 at these parameter values; L_0 tells how frequently we can tolerate interference with a satisfactory process and L_1 says how long we can go before a deterioration to μ_1 should be noticed. Usually L_0 will be quite large, of the order of 250 or 1000, while L_1 will be small, less than about 10.

Graphs of the A.R.L. for schemes for controlling the mean of a normal distribution against two-sided deviations have been constructed by Goldsmith and Whitfield (6) and tables and a nomogram have been prepared for the one-sided case by Ewan and Kemp (5). In the latter paper the authors use approximate solutions of the integral equations arising from the defining Wald test (7, 11), have derived an empirical approximation to the A.R.L. at μ_1 and have shown that instead of plotting the cumulative sum of deviations from the target value, $\sum (x_i - \mu_0)$, it is better to plot $\sum (x_i - k)$ where k is a reference value. If k is taken to be a convenient quantity near to $(\mu_0 + \mu_1)/2$ the A.R.L. at μ_0 is near its maximum for a given A.R.L. at μ_1 . The empirical approximation for the A.R.L. at μ_1 when the observations have unit variance is

$$L_1 = 1 + \frac{h}{\mu_1 - k}$$

so that both defining characteristics of the one-sided cumulative sum scheme, k and h, can be determined approximately from μ_0 , μ_1 and L_1 .

When we are dealing with other distributions, for example, controlling the S. D. of a normal population using the range or mean range of samples, the above approximation for the A.R.L. at poor quality may be used and at good quality the tabled results for the normal mean will be a satisfactory approximation. Naturally in these applications the μ's will refer to the mean of the distribution of the statistic concerned and the S. D. of this distribution will also be required; in the example above, we would use the mean and S. D. of the range distribution as given, e.g. in Pearson's tables.

Tables of schemes for controlling the fraction defective of a process calculated using the Wald test exact results (3, 8) were given in (9) —the A.R.L. in these

tables is the average number of observations and *not* samples. Detailed examples of the choice of schemes based on the above principles are given in the papers cited.

Practical Details

Since such emphasis is placed on the visual impact of these schemes some attention needs to be paid to the scale on which the charts are plotted. We are looking for a change in the direction of the mean path of a cumulative sum and some experiments by Barnard have shown that changes in mean of moderate amount are most easily noticed when one horizontal step is about equal to a 2σ step vertically. A similar conclusion has been reached independently by Truax (17).

The above applications of cumulative sum charts to process control follow those of the Shewhart chart quite closely and in them they are more precise and sensitive in operation and more easily interpreted visually. The other most important requirements of the charts kept in a process inspection scheme is to the post mortem examination of results. This phase is usually concerned with test and estimation problems; did a process change occur in the time covered by the given records, and if so, what kind of change, how much and when (13,14)? Again, cumulative sum procedures compare well with Shewhart ones. The point of change is shown much more clearly on a sum chart by a change in the direction of the mean path than it is on the classical chart by a slight difference in the mean level of a set of points.

The Run Length Distribution

In the case of one-sided schemes it has been shown (9, 5) that the distribution of the run length will be approximately geometric when the A.R.L. is large. The scheme is equivalent to the repetition of Wald tests; the number of repetitions of the tests has a geometric distribution and when the A.R.L. is large this dominates the form of the distribution of the number of samples. Page showed that the distribution of the run length, l, is approximated for large n by

$$pr(l \leq n) \doteqdot 1 - P^{v+1}$$

where $v = n/N_1$ and where P, N_1 , are respectively the O.C. and A.S.N. conditional upon acceptance of the defining Wald test. If we use the result that the A.R.L., L, is given by

$$L = N/(1 - P)$$

where N is the unconditional A.S.N. and put $N_1 \doteqdot N$ we obtain Kemp's result

$$p(l \leq n) \doteqdot 1 - e^{-n/L}.$$

Thus for the one-sided case at least and for n, L large, the run length distribution is specified completely by the A.R.L. and so we have good justification for using the A.R.L. to choose our schemes.

Further Problems

The cumulative sum charts can be used and are being used as they stand; but they have already suggested several areas in which further work would be rewarding. The problems of estimating the position and the size of changes are worth more attention. The effects of correlated observations have been examined in one case by Goldsmith and Whitfield; wider study is desirable. The problems of inspection procedures are related to the whole field of control methods and we need to know how action based upon cumulative sum and Shewhart charts affects the process. Different criteria of choice of scheme will be relevant and different rules for indicating action and its amount may be appropriate; rules with curved boundaries may be most satisfactory. There will be smaller problems arising from minor modifications to the inspection schemes; can a scheme for controlling both mean and S. D. on the same chart be constructed as is possible on a warning line chart (12); can other quick methods for example, gauging (16) be applied to the inspection schemes which lose little in efficiency and give much in convenience of application?

References

1. Aroian, L. A. & Levene, H. (1950), *J. Amer. Statist. Assoc. 45*, 520.
2. Barnard, G. A. (1959), *J. Roy. Statist. Soc. B. 21*, 239.
3. Burman, J. P. (1946), *J. Roy. Statist. Soc. Suppl. 8*, 98.
4. Duncan, A. J. (1956), *J. Amer. Statist. Assoc. 51*, 228.
5. Ewan, W. D. & Kemp, G. D. (1960), *Biometrika 47*, 363.
6. Goldsmith P. L. & Whitfield H. (1961), *Technometrics 3*, 1.
7. Kemp, K. W. (1958), *J. Roy. Statist. Soc. B, 20*, 379.
8. Moore, P. G. (1958), *Biometrika 45*, 89.
9. Page, E. S. (1954), *Biometrika, 41*, 101.
10. Page, E. S. (1954), *J. Roy. Statist. Soc. B, 16*, 131.
11. Page, E. S. (1954), *J. Roy. Statist. Soc. B, 16*, 136.
12. Page, E. S. (1955), *Biometrika 42*, 243.
13. Page, E. S. (1955), *Biometrika 42*, 523.
14. Page, E. S. (1957), *Biometrika 44*, 249.
15. Shewhart, W. A. (1931), The Economic Control of the Quality of Manufactured Product. Macmillan, New York.
16. Stevens, W. L. (1948), *J. Roy. Statist. Soc. B, 10*, 54.
17. Truax, H. Mack. (1960), Report on Cumulative Sum Charts, Atlas Powder Co.
18. Walker, A. M. (1950), *J. Roy. Statist. Soc. B, 12*, 301.
19. Weiler, H. (1952), *Ann. Math. Statist. 23*, 247.
20. Weiler, H. (1953), *J. Amer. Statist. Assoc. 48*, 816.
21. Weiler, H. (1954), *J. Amer. Statist. Assoc. 49*, 298.

Average Run Lengths in Cumulative Chart Quality Control Schemes

P. L. GOLDSMITH AND H. WHITFIELD

Imperial College, London

Average run lengths are evaluated for V-mask quality control schemes based on cumulative deviation charts when the observations are Normally distributed and either independent or members of a certain serially correlated class.

1. CUMULATIVE CHART QUALITY CONTROL SCHEMES

The Shewhart quality control chart with fixed control lines suffers from the disadvantage that the observations are viewed independently, no account being taken of runs of observations all higher or all lower than the long-term mean of the process. This results in a relative insensitivity to moderate changes in the short-term mean value. For example consider a chart with fixed limit lines at $\pm$ 3.09σ controlling the mean of a Normal population of constant variance, σ^2: there is an average run length of about 55 single observations before a shift in current mean of one standard deviation from the long-term mean is detected.

This situation can be improved by having warning lines at $\pm$ 1.96σ in addition to the action lines at $\pm$ 3.09σ. Further schemes have been devised with alternative positions for the control lines, some taking account of runs of observations above (or below) the long-term mean. For example Page (4) considers rules of the type "take action if any point falls outside the action lines or if any m out of the last n pointsf all outside the warning lines." However such rules are inefficient compared with control schemes based on cumulative charts. Page (3) assigns a score x_j to the jth observation and plots a cumulative score,

$$S_n = \sum_{i=1}^{n} x_i ,$$

against n. The system of scoring is chosen so that the mean path on the chart is downwards when quality is satisfactory and upwards when quality is unsatisfactory. For a one-sided control scheme, Page adopts the rule "take action after the nth observation if $S_n - \text{Min.}_{0 \le i < n} (S_i) \ge h$", and for detecting changes in either direction suggests "take action after the nth observation if either $S_n - \text{Min.}_{0 \le i < n} (S_i) \ge h$ or $\text{Max.}_{0 \le i < n} (S_i) - S_n \ge h'$" where the decision intervals, h and h', are positive constants. He discusses the calculation of average run lengths for both of these schemes. A useful amendment to these rules has been proposed by Ewan and Kemp (2), who form cumulative sums of the differences of observations from a reference level mid-way between the acceptable and rejectable quality levels. These values are plotted only when the

cumulative difference is relevant towards taking a decision that quality has changed adversely.

Following essentially the same procedure, Barnard (1) takes

$$S_n = \sum_{j=1}^{n} x_j$$

for the cumulative score where x_j is the algebraic difference between the jth observation and the long-term process mean, and the corresponding chart is perhaps best termed a 'cumulative deviation chart'. From this definition, it follows that changes in process mean are reflected by proportionate changes in the average slope of the cumulative graph, the mean path being horizontal when the process is on target. Trends in the process mean which are not easily discernible on a Shewhart chart show up remarkably clearly on a cumulative deviation chart, though a certain amount of care is needed to avoid being misled by trends which are only apparent.

For a two-sided control scheme based on this chart, a symmetric V-shaped mask is advocated with the vertex of the V pointing horizontally forwards and at a distance d ahead of the current point (n, S_n). This is shown in Figure 1:

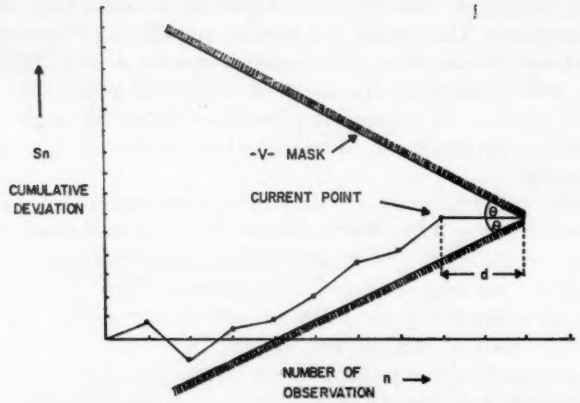


FIGURE 1—V-mask superimposed on cumulative deviation chart.

θ is the angle between each of the limbs of the V and the horizontal. The V can be etched on a perspex sheet and the sheet moved into the correct position as each value is plotted on the cumulative chart. A lack of process control is indicated when a previously plotted point falls outside the limbs of the V-mask: the origin counts as a previous point for this purpose ($S_0 = 0$). The graph is recommenced or an appropriate correction is made when the process mean is reset after an out-of-control warning. By virtue of their simplicity it is very straight-forward to operate such schemes in the chart-room of a manufacturing plant, and indeed the plotting may be carried out on a suitably adapted automatic recorder.

Page (5) has derived a similar test procedure with $d = 0$, assuming that the size of the change in the process mean (if it occurred) was known.

The behaviour of the Barnard control scheme depends on the two parameters d, θ. One way of selecting the proper parameter values for a particular application is to try out a variety of masks on historical records. The shape of the mask is adjusted so that past changes in the current mean are deduced in a reasonable time, whilst avoiding the alternative trouble of giving indications of lack of control when the fluctuations are really due to chance error. However, this method is subjective, and it may be preferable to decide on two desired average run lengths:

L_0 , when the process is on target, and
L_1 , when the current mean is $k\sigma$ off target,

and to determine d and θ accordingly: σ is the standard error of the sample means, assumed known and constant. If the rate at which samples are taken is constant, the average run length is a measure of how much of the output is of suspect quality, and it may be used as one simple criterion for comparing different schemes. Hence average run lengths, L, have been calculated for various values of d and θ with k in the range $0 - 3$. These values will enable the statistician to assess the behaviour of a V-mask he plans to use and will facilitate the design of suitable control schemes.

This paper does not discuss directly the important use of a cumulative control chart as a method for estimating the current mean of the process. This may be achieved by superimposing a parabolic or quartic mask on the chart; the vertex of the mask is placed over the current point and the axis is rotated so as to include the maximum number of consecutive past points within the arc of the mask. This gives greater weight to the more recent observations, and the final direction of the mask axis corresponds to the estimated current mean (1).

2. Average Run Lengths from Independent Normal Observations

The choice of θ will also depend on the scale used when plotting the cumulative deviation chart. For purposes of standardisation, it has been assumed that the plotting interval on the horizontal axis is equal to 2σ on the vertical axis, so that when the current process mean shifts 2σ from the target, the mean path of the cumulative graph will make an angle of 45° with the horizontal. This arrangement permits a rapid visual appreciation of the behaviour of the process from the appearance of the cumulative graph. When the plotting interval on the horizontal axis is equal to $\nu\sigma(\nu \neq 2)$ on the vertical axis then the values of tan θ given below should be multiplied by $2/\nu$.

It is assumed throughout that the observations (means of samples) come from a Normal population $N(k\sigma, \sigma^2)$; it is further assumed in this section that the observations are mutually independent.

Figures 2, 3, 4 and 5 graph average run lengths, L, (on a logarithmic scale) against the displacement of the current mean for masks with lead distances of 1, 2, 5 and 8 units respectively, where the unit is the plotting interval on the horizontal axis. The values of tan θ have been chosen so as to give a range of L_0's from 20 to 1500, and the graphs may be interpolated. It will be seen that certain combinations of d and θ have similar characteristics. The evaluation

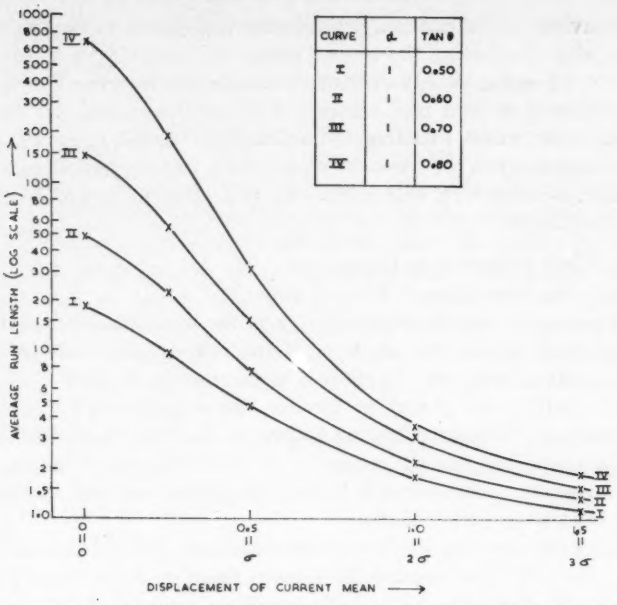


FIGURE 2—Average run length v. Displacement of current mean for symmetric V-masks with $d = 1$.

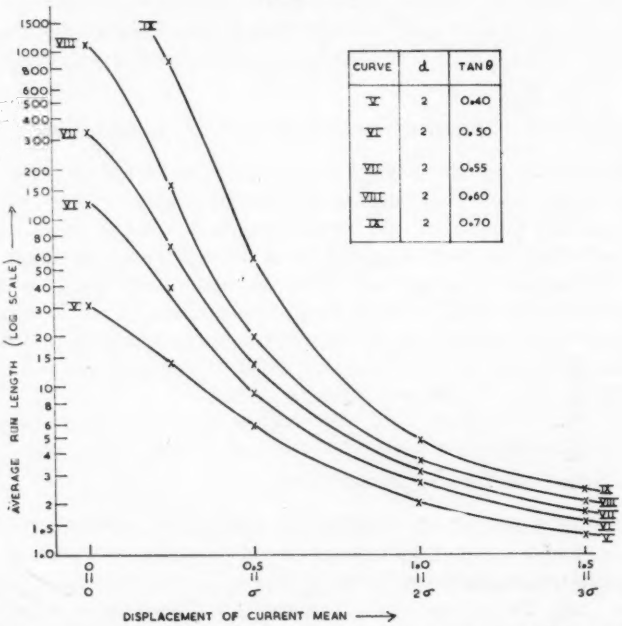

CURVE | d | TAN θ
V | 2 | 0.40
VI | 2 | 0.50
VII | 2 | 0.55
VIII | 2 | 0.60
IX | 2 | 0.70
AVERAGE RUN LENGTH (LOG SCALE)
DISPLACEMENT OF CURRENT MEAN
0 0.5 1.0 1.5
0 σ 2σ 3σ

FIGURE 3—Average run length v. Displacement of current mean for symmetric V-masks with $d = 2$.

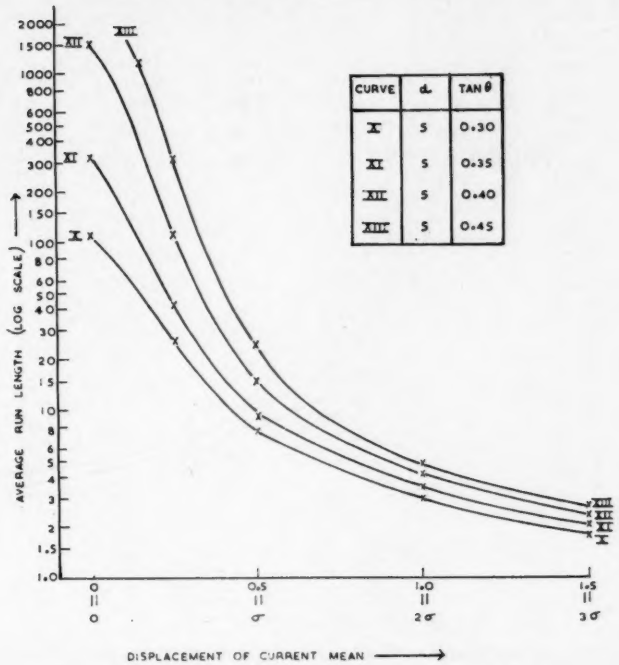


FIGURE 4—Average run length v. Displacement of current mean for symmetric V-masks with $d = 5$.

was carried out by a Monte Carlo simulation on a Ferranti 'Mercury' digital computer: a sequence of pseudo-random Normal deviates was generated by a fast table look-up procedure. Each calculated L has a coefficient of variation less than 10 per cent, an accuracy which is regarded as adequate for practical purposes.

The results may also be expressed in terms of two empirical formulae as follows:

$$\log_{10} \log_{10} L_0 = -0.5244 + 0.0398\, d + 1.1687 \tan\theta + 1.2641 \tan\theta \log_{10} d$$

$$L_1 = \frac{2\, d \tan\theta}{k - 2 \tan\theta} + \frac{2}{3}$$

The first of these equations provides a good approximation for L_0 at all the values of d and θ investigated, and the second evaluates L_1 adequately when $1.5 < k < 4$.

Suppose, for example, we require $L_0 = 500$ when the process is on target and $L_1 = 15$ when $k = 1$. Then the mask with parameters $d = 2$, $\tan\theta = 0.565$ is satisfactory. It is interesting to note that the corresponding standard Shewhart chart with control lines 3.09σ either side of the long-term mean has $L_0 = 500$ but $L_1 = 55$, exemplifying the power of the cumulative chart scheme.

A comparison of average run lengths between standard Shewhart and cumu-

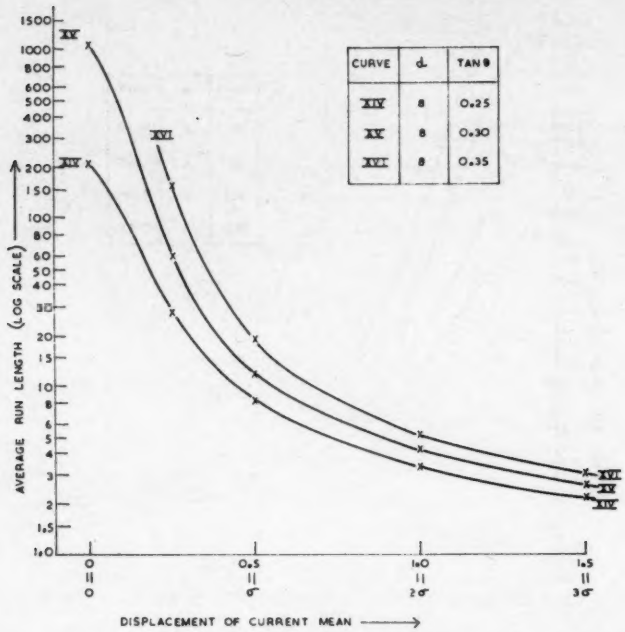


FIGURE 5—Average run length v. Displacement of current mean for symmetric V-masks with $d = 8$.

lative deviation schemes is shown in Figure 6 for two typical V-masks, the width of the Shewhart control lines being chosen so as to give approximately equal values for L_0 in each case. The data are listed in Tables 1 and 2.

The power of the cumulative scheme is most evident for moderate deviations of the current mean: a reduction factor of 4 in the average run length is possible with the first of these two masks. Exceptionally large changes are picked up a little more quickly by the Shewhart chart. It may sometimes be better to plot a Shewhart chart and a cumulative deviation chart simultaneously to obtain a picture of the process variation.

The values of L may also be used to derive the average run lengths of the two-sided scheme proposed by Ewan and Kemp (2). For it can be shown that

TABLE 1

$d = 5$, $\tan \theta = 0.35$ *Equivalent control lines at* $\pm 2.96\sigma$

Average run length with	Deviation of current mean						
	0	$\frac{1}{2}\sigma$	σ	2σ	3σ	4σ	5σ
Cumulative Scheme	319	42.3	9.23	3.50	2.09	1.59	1.21
Shewhart Scheme	320	137	39.5	5.89	1.92	1.17	1.02

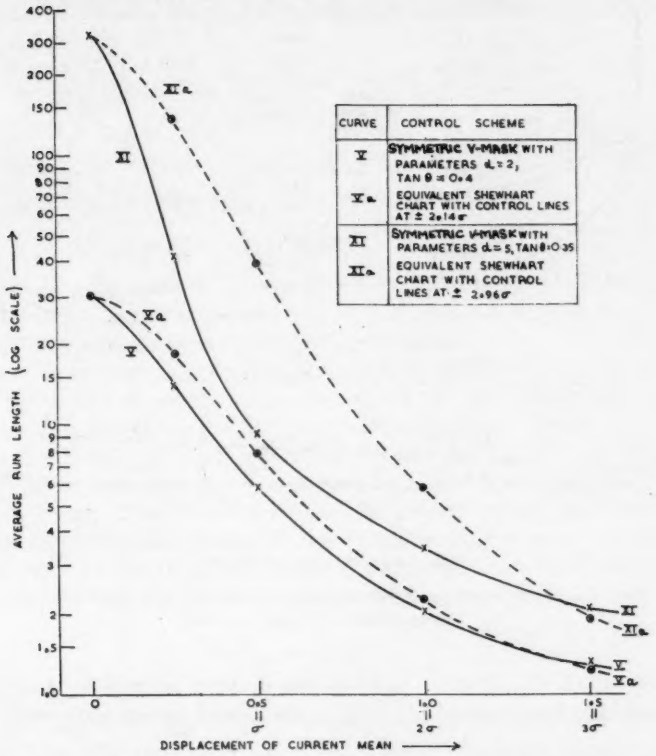


FIGURE 6—Comparison of average run lengths between Shewhart and cumulative schemes.

the behaviour of a symmetric V-mask scheme is statistically equivalent to a pair of adjoint decision interval plots. If points on the latter charts are derived from the means of samples of size N having within sample standard deviation σ', then the relationship between the parameters of the two schemes is given by the equations

$$\Theta = \frac{2\sigma'}{\sqrt{N}} \tan \theta \qquad h = \frac{2\sigma'}{\sqrt{N}} d \tan \theta,$$

TABLE 2

$d = 2$, $\tan \theta = 0.40$ *Equivalent control lines at* $\pm 2.14\sigma$

Average run length with	Deviation of current mean						
	0	$\frac{1}{2}\sigma$	σ	2σ	3σ	4σ	5σ
Cumulative Scheme	30.7	13.9	5.84	2.04	1.30	1.05	1.005
Shewhart Scheme	30.5	18.1	7.75	2.24	1.24	1.03	1.002

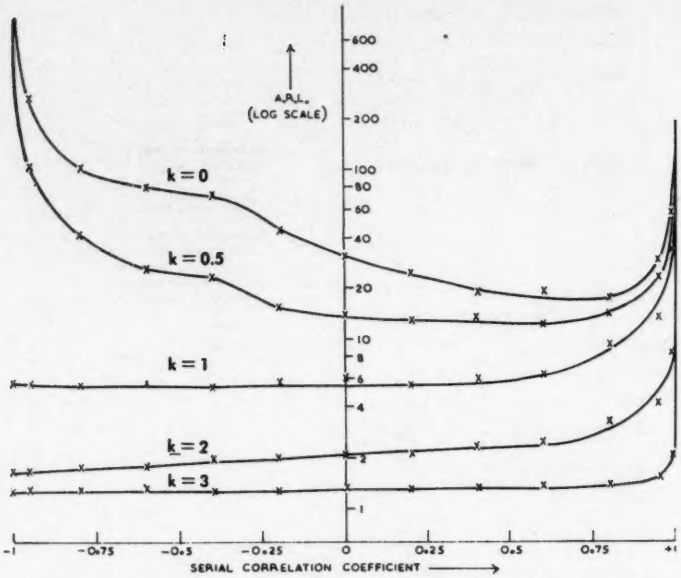


FIGURE 7—Average run length v. Serial correlation coefficient for symmetric V-masks with parameters $d = 2$, $\tan \theta = 0.4$

where $\Theta = \frac{1}{2}$(A.Q.L.-R.Q.L.) and h is the decision interval in the Ewan and Kemp scheme. The average run lengths are related by the reciprocal rule:

$$\frac{1}{L} = \frac{1}{L_U} + \frac{1}{L_L},$$

where L is the A.R.L. of the V-mask scheme, L_U is the A.R.L. of the upper decision interval plot, and L_L is the A.R.L. of the lower decision interval plot.

3. Effect of Serial Correlation Between the Observations

In practice, industrial plant observations are often serially correlated, so that it is important to know how serial correlation effects the values of L. Correlated pseudo-random deviates x_i, from $N(k\sigma, \sigma^2)$ were therefore generated according to the first order autoregressive sequence:

$$y_i = ay_{i-1} + (1 - a^2)^{\frac{1}{2}}\epsilon_i \qquad j = 1, 2, \cdots,$$
$$x_i = k\sigma + y_i$$

where ϵ_i $(j = 1, 2, \cdots)$ and y_0 are pseduo-random Normal deviates from $N(0, \sigma^2)$ and $|a| \leq 1$. Then

$$\mathcal{C}(y_r, y_{r-s}) = \sigma^2 a^s$$

and $$s = 0, 1, 2, \cdots, r.$$

$$\rho(y_r, y_{r-s}) = a^s$$

When $a = 0$ we have the case of independent observations considered above. When $a = 1$,

$$x_j = k\sigma + y_0 \quad \text{for all} \quad j,$$

$$S_n = n(k\sigma + y_0),$$

and it follows that the average run length is infinite. When $a = -1$,

$$x_{2j-1} = k\sigma - y_0\,, \qquad x_{2j} = k\sigma + y_0\,,$$

$$S_{2n-1} = (2n-1)k\sigma - y_0\,, \qquad S_{2n} = 2nk\sigma:$$

this leads to $L = \infty$ when the deviation of the current mean is less than or equal to $\tan\theta$, but for larger deviations the finite average run lengths can be calculated from tables of the Normal probability integral.

Figures 7 and 8 graph the average run length against the coefficient a for various deviations of current mean: Figure 7 is for the V-mask with lead distance $d = 2$ units and $\tan\theta = 0.4$ (on the $\sigma = \frac{1}{2}$ scale as above), and Figure 8 refers to the V-mask with the same lead distance but $\tan\theta = 0.5$. There is a discontinuity in the graphs at $a = -1$ when $k\sigma > \tan\theta$ and $d\tan\theta/(k\sigma - \tan\theta)$ is an even integer. It will be seen that for large deviations of the current mean the serial correlation has little effect on L, except near $a = 1$. But when k is small, positive correlation in the process increases the effectiveness of the cumulative chart control scheme (halving the average run length is some cases),

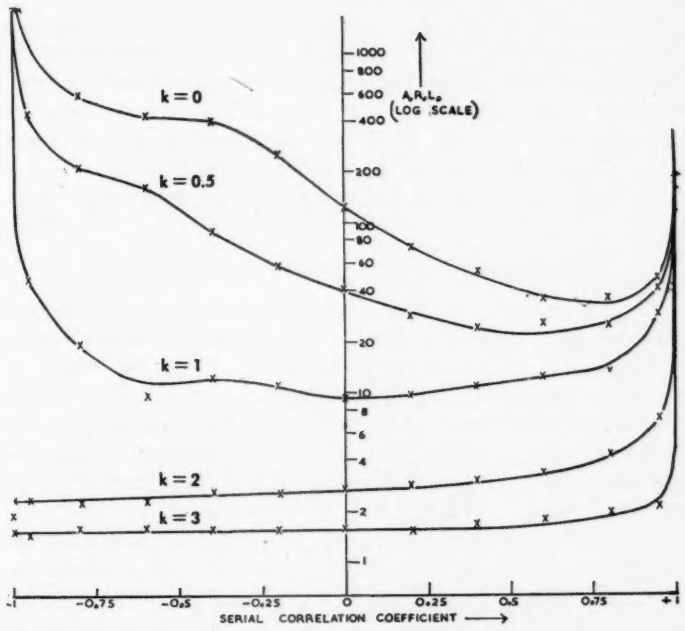


FIGURE 8—Average run length v. Serial correlation coefficient for symmetric V-mask with parameters $d = 2$, $\tan\theta = 0.5$.

while negative correlation leads to higher average run lengths. These curves may be regarded as typical.

When the value of a is known or can be estimated for some process it is desired to control, the corrected values of the average run length should be used in deciding on the appropriate values of c and θ. In the case of positive correlation this will avoid troubles due to too frequent an interruption into a satisfactory process.

The authors acknowledge with thanks discussions on this paper with Professor G. A. Barnard of Imperial College, London.

References

(1) Barnard, G. A., 'Control Charts and Stochastic Processes', *J. Roy. Stat. Soc. B*, Vol. 21 (1959) pp. 239–271.
(2) Ewan, W. D. and Kemp, K. W., 'Sampling Inspection of Continuous Processes with no Autocorrelation Between Successive Results', *Biometrika*, Vol. 47 (1960) pp. 363–380.
(3) Page, E. S., 'Continuous Inspection Schemes', *Biometrika*, Vol. 41 (1954) pp. 100–114.
(4) Page, E. S., 'Control Charts with Warning Lines', *Biometrika*, Vol. 42 (1955) pp. 243–257.
(5) Page, E. S., 'On Problems in which a Change in a Parameter Occurs at an Unknown Point', *Biometrika*, Vol. 44 (1957) pp. 248–252.

Prediction Regions for Several Predictions from a Single Regression Line*

GERALD J. LIEBERMAN

Stanford University

When a linear relationship has been fitted by least squares, the methods for securing a prediction interval for the response at some fixed value of the independent variable are explained in many statistical text books. This paper describes the somewhat more complex problem of determining the joint prediction interval for the responses at each of K separate settings of the independent variables when all K predictions must be based upon the original fitted model.

I. INTRODUCTION

Determining the relationship between several variables is a problem which often arises in engineering. In particular, a relationship of the form

$$\text{Average value of } y = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p$$

is usually assumed and estimates of β_0 , β_1 , $\cdots$, β_p are obtained from sample data. More important, however, these results are used to make statements about future values of y for given values of the X's. For example, consider the following simple situation. The speed of a missile is a critical factor, which is determined only by measurement after firing. On the other hand, the orifice opening of the valve which admits the fuel is easily obtained by bench tests. Suppose that there exists an underlying relationship between these two variables of the form

$$\text{Average value of speed} = \beta_0 + \beta_1 \text{ (orifice opening)}.$$

Hence, speed corresponds to y and orifice opening corresponds to X_1 ($p = 1$) in the above linear relationship. N missiles are produced and fired so that N pairs of observations, (X_{11}, y_1), (X_{12}, y_2), $\cdots$, (X_{1N}, y_N) are taken. Estimates of β_0 and β_1 are obtained, usually by the method of least squares. If a new missile is produced a prediction of its speed based on a knowledge of its orifice opening may be desirable. The usual techniques can be used to obtain a prediction interval such that the speed of the missile, having a given orifice opening X^*, will lie in the interval with preassigned probability, say 0.95. Expressions for such an interval can be obtained from any standard text.[1]

* Work done under the sponsorship of the Office of Naval Research under Contract N6 onr-25126.

[1] For example, see Chapter 9 of *Engineering Statistics* by Bowker and Lieberman, Prentice Hall, 1959.

A more difficult problem arises when K additional missiles are produced (rather than one), and a prediction region is desired for the K speeds given the K orifice openings and based on the original N pairs of observations. Using the results for the one missle K times is incorrect since the prediction intervals are *not* independent. They are all based upon the same original N observations. It is the purpose of this paper to give three methods for obtaining a prediction region for K future observations $(y_1^*, y_2^*, \cdots, y_K^*)$ based upon the same estimated linear regression given the values of the independent variables

$$(X_{11}^*, X_{21}^*, \cdots, X_{p1}^*; X_{12}^*, X_{22}^*, \cdots, X_{p2}^*; \cdots; X_{1K}^*, X_{2K}^*, \cdots, X_{pK}^*).$$

II. An Exact Prediction Region

(1) Let y_α, $\alpha = 1, \cdots, N$ be independent normally distributed random variables with

$$E(y_\alpha) = \beta_0 + \beta_1 X_{1\alpha} + \cdots + \beta_p X_{p\alpha}$$

and

$$\sigma^2(y_\alpha) = E[y_\alpha - E(y_\alpha)]^2 = \sigma^2 \quad \text{for all} \quad \alpha.$$

Define

$$a_{ij} = \sum_{\alpha=1}^{N} X_{i\alpha}X_{j\alpha} - N\bar{X}_i\bar{X}_j; \qquad A = (a_{ij}).$$

Let $b_0, b_1, \cdots, b_p$ be the least square estimates of $\beta_0, \beta_1, \beta_2, \cdots, \beta_p$, respectively. Let

$$X_r^* = (X_{r1}^*, X_{r2}^*, \cdots, X_{rp}^*), \qquad r = 1, 2, \cdots, K$$

be a fixed value of the vector

$$X = (X_1, X_2, \cdots, X_p),$$

and

$$\hat{y}_r^* = b_0 + b_1X_{r1}^* + b_2X_{r2}^* + \cdots + b_pX_{rp}^*.$$

Then the covariance of $\hat{y}_r^*$ and $\hat{y}_s^*$ is given by

$$(2) \qquad \sigma_{\hat{y}_r^*\hat{y}_s^*} = \sigma^2\left[\frac{1}{N} + \sum_{i=1}^{p}\sum_{j=1}^{p} a^{ij}(X_{ri}^* - \bar{X}_i)(X_{sj}^* - \bar{X}_j)\right]$$

where a^{ij} is the element in the ith row and jth column in the inverse matrix of A.

Denote by y_r^* the future rth observation on y corresponding to the vector X_r^*. The quantities $y_r^* - \hat{y}_r^*$ are normally distributed random variables with zero expectation and covariance

$$(3) \qquad \begin{aligned} E(y_r^* - \hat{y}_r^*)(y_s^* - \hat{y}_s^*) &= E[\hat{y}_r^* - E(y_r^*)][\hat{y}_s^* - E(y_s^*)], \qquad r \neq s \\ E(y_r^* - \hat{y}_r^*)^2 &= E[y_r^* - E(y_r^*)]^2 + E(\hat{y}_r^* - E(y_r^*)]^2, \qquad r = s. \end{aligned}$$

Let $\sum = (\sigma_{rs})$ be the matrix of covariance and variances for the random vari-

ables $(y^* - \hat{y}^*)$. From (1), (2) and (3) it follows that

$$\sigma_{rr} = \sigma^2\left[1 + \frac{1}{N} + \sum_{i=1}^{p}\sum_{j=1}^{p} a^{ij}(X_{ri}^* - \bar{X}_i)(X_{rj}^* - \bar{X}_j)\right]$$

(4)

$$\sigma_{rs} = \sigma^2\left[\frac{1}{N} + \sum_{i=1}^{p}\sum_{j=1}^{p} a^{ij}(X_{ri}^* - \bar{X}_i)(X_{sj}^* - \bar{X}_j)\right] \quad \text{for} \quad r \neq s.$$

Thus $(y^* - \hat{y}^*)$ has a multivariate normal distribution with zero mean and covariance matrix $\sum$. Let

$$Q = \sum_{r=1}^{K}\sum_{s=1}^{K} \sigma^{rs}(y_r^* - \hat{y}_r^*)(y_s^* - \hat{y}_s^*)$$

where σ^{rs} is the element in the rth row and sth column in the inverse matrix of $\sum$. Then Q has a χ^2 distribution with K degrees of freedom. If we define

$$s_{y|X}^2 = \sum_{\alpha=1}^{N} [y_\alpha - (b_0 + b_1X_{1\alpha} + \cdots + b_pX_{p\alpha})]^2/(N - p - 1)$$

then $(N - p - 1)\, s_{y|X}^2/\sigma^2$ is independent of Q and has a χ^2 distribution with $N - p - 1$ degrees of freedom. It follows, therefore, that the quantity

$$U = \frac{\sum_{r=1}^{K}\sum_{s=1}^{K} \sigma^{rs}(y_r^* - \hat{y}_r^*)(y_s^* - \hat{y}_s^*)}{Ks_{y|X}^2/\sigma^2}$$

has an F distribution with K and $N - p - 1$ degrees of freedom and is independent of σ^2. Thus for any level of significance α the equation

$$U \leq F_{\alpha;K,N-p-1}$$

gives a prediction region for $y^* = (y_1^*, \cdots, y_K^*)$, where $F_{\alpha;K,N-p-1}$ is the α percentage point of the F distribution having K and $N - p - 1$ degrees of freedom. Whether a particular set of values $y_1^*, y_2^*, \cdots, y_K^*$ falls within the prediction region is easily ascertained by substituting these values into the expression for U and determining whether the result is less than or equal to $F_{\alpha;K,N-p-1}$.

III. Approximate Prediction Regions

The prediction region given in the previous section is clearly an ellipsoid. Although this ellipsoid yields an exact prediction region, its interpretation may be difficult. Often, one is interested in prediction regions which are intervals for each future observation. This section describes two such regions. Both of these regions will be approximate in that the probability will be *at least* $(1 - \alpha)$ that they contain the future observations $y_1^*, y_2^*, \cdots, y_K^*$.

Following the notation given in Section II, an exact prediction interval for $K = 1$ is given by

$$y_1^* \pm (t_{\alpha/2;N-p-1})s_{y|X}\sqrt{1 + \frac{1}{N} + \sum_{i=1}^{p}\sum_{j=1}^{p} a^{ij}(X_{1i}^* - \bar{X}_1)(X_{1j}^* - \bar{X}_j)}$$

where $t_{\alpha/2;N-p-1}$ is the $\alpha/2$ percentage point of the t distribution having $N - p - 1$ degrees of freedom. Suppose there are now K future observations and the above prediction interval is used for each (varying the value of the independent variables each time of course) replacing $t_{\alpha/2;N-p-1}$ by $t_{\alpha/2K;N-p-1}$. The probability that *all* of the K future observations fall into their respective intervals is *at least* $(1 - \alpha)$. This is easily seen by taking the special case of $K = 2$. Let A be the event that the first future observation falls outside of its prediction interval, and let B be the event that the second future observation falls outside of its prediction interval. Then A or B is the failure of the pair of intervals to simultaneously bracket the two future observations. Choose $P(A) = P(B) = \alpha/2$. Hence

$$P(A \quad \text{or} \quad B) = P(A) + P(B) - P(AB) \leq P(A) + P(B) = \alpha.$$

It then follows that the probability that both future observations simultaneously fall into their prediction intervals is at least $1 - \alpha$.

The second prediction region is obtained by circumscribing a rectangular region about the ellipsoid. Since this rectangular region includes the ellipsoid, the probability will be at least $(1 - \alpha)$ that it contains the future observations $y_1^*, y_2^*, \cdots, y_K^*$. The values of the end points of the intervals which describe the rectangular region are easily obtained. Scheffé[2] has pointed out that an ellipsoid can be generated by tracing out all of the supporting hyperplanes. In particular, the region of interest can be described by a small subset of these hyperplanes; namely, those which occur at the maximum and minimum of the ellipsoid. These supporting hyperplanes lead to the following intervals for the future observations

$$\hat{y}_i^* \pm s_{y|x}\sqrt{KF_{\alpha;K,N-p-1}\sigma_{ii}/\sigma^2}; \qquad i = 1, 2, \cdots, K.$$

This set of intervals generates the rectangular region which circumscribes the ellipsoid.

IV. Example

Consider the example mentioned earlier about predicting the speed of a missile from a measurement of its nozzle opening. Suppose 5 pair of observations are taken and are given below

speed (miles/hr)	5,940	5,550	5,070	4,520	4,360
opening (inches)	1.40	1.36	1.34	1.32	1.31.

If two more missiles are produced having nozzle openings of 1.37 and 1.33 inches respectively, 90% prediction regions using the three methods given in this paper are as follows:

Exact Method.

The least squares estimates of β_0 and β_1 are given by

---

[2] For example, see The Analysis of Variance by H. Scheffé, John Wiley, 1959.

$$b_1 = \frac{\sum_{i=1}^{5}(X_i - \bar{X})(y_i - \bar{y})}{\sum_{i=1}^{5}(X_i - \bar{X})^2} = 18273.44$$

and

$$b_0 = \bar{y} - b_1\bar{X} = -19508.05$$

so that the equation of the estimated line is

$$\hat{y} = -19508.05 + 18273.44X.$$

Furthermore

$$s_{y|x} = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n-2}} = 165.95.$$

The variance and covariance matrix $\sum$ is given by

$$\sum = \begin{bmatrix} \sigma_{11} & \sigma_{21} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}$$

where

$$\sigma_{11} = \sigma^2\left[1 + \frac{1}{N} + \frac{(X_1^* - \bar{X})^2}{\sum (X_i - \bar{X})^2}\right] = 1.31\sigma^2$$

$$\sigma_{12} = \sigma_{21} = \sigma^2\left[\frac{1}{N} + \frac{(X_1^* - \bar{X})(X_2^* - \bar{X})}{\sum (X_i - \bar{X})^2}\right] = .275\sigma^2$$

and

$$\sigma_{22} = \sigma^2\left[1 + \frac{1}{N} + \frac{(X_2^* - \bar{X})^2}{\sum (X_i - \bar{X})^2}\right] = 1.25\sigma^2.$$

Hence

$$\sum^{-1} = \frac{1}{\sigma^2}\begin{bmatrix} 0.799 & -0.176 \\ -0.176 & 0.839 \end{bmatrix}$$

so that

$$\sigma^{11} = \frac{0.799}{\sigma^2}, \qquad \sigma^{12} = \sigma^{21} = \frac{-0.176}{\sigma^2}, \qquad \sigma^{22} = \frac{0.839}{\sigma^2}.$$

The prediction region is then given by

$$U = \frac{\sigma^{11}(y_1^* - \hat{y}_1^*)^2 + 2\sigma^{12}(y_1^* - \hat{y}_1^*)(y_2^* - \hat{y}_2^*) + \sigma^{22}(y_2^* - \hat{y}_2^*)^2}{2s_{y|x}^2/\sigma^2} \leq F_{0.10;2,3} = 5.46$$

or

$$U = 0.0000145y_1^{*2} + 0.0000152y_2^{*2} - 0.00000638y_1^*y_2^* - 0.130y_1^* - 0.111y_2^* + 623.996 \leq 5.46.$$

This ellipse is plotted in Figure 1.

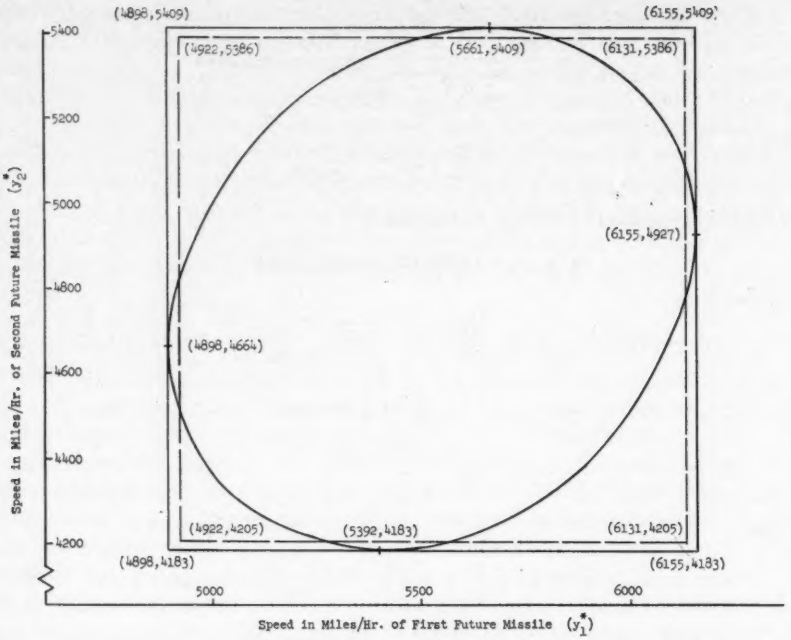


FIGURE 1—Prediction regions for two future missiles having nozzle openings of 1.37 and 1.33 inches respectively.

Approximate Methods.

For the first approximate method the prediction interval for y_1^* is given by

$$\hat{y}_1^* \pm (t_{0.025,3})s_{y|x}\sqrt{1 + \frac{1}{5} + \frac{(X_1^* - \bar{X})^2}{\sum (X_i - \bar{X})^2}}$$

or

$$[4922,\ 6131].$$

The prediction interval for y_2^* is given by

$$\hat{y}_2^* \pm (t_{0.025,3})s_{y|x}\sqrt{1 + \frac{1}{5} + \frac{(X_2^* - \bar{X})^2}{\sum (X_i - \bar{X})^2}}$$

or

$$[4205,\ 5386].$$

The probability that both future observations fall within their respective intervals is greater than 0.90. This region is shown by the dashed inner rectangle in Figure 1.

For the second approximate method (prediction region which circumscribes the ellipsoid), the prediction interval for y_1^* is given by

$$\hat{y}_1^* \pm s_{y|x}\sqrt{2F_{0.10;2,3}\sigma_{11}/\sigma^2} = 5526.56 \pm 165.95\sqrt{(2)(5.46)(1.31)}$$

or

$$[4898,\ 6155].$$

The prediction interval for y_2^* is given by

$$\hat{y}_2^* \pm s_{y|x}\sqrt{2F_{0.10;2,3}\sigma_{22}/\sigma^2} = 4795.63 \pm 165.95\sqrt{(2)(5.46)(1.25)}$$

or

$$[4183,\ 5409].$$

The probability that both future observations fall within their respective intervals is greater than 0.90. This region is shown by the solid outer rectangle in Figure 1.

V. Remarks

It is rather surprising that the exact prediction region given by the ellipse does not lie totally within the rectangle obtained from using the first approximate method. This approximation appears to be crude. N is small (five) in this example so that the prediction for the future observations is strongly dependent upon the estimate of the initial regression line. If N were large, this approximate method would be expected to give good results since the regression line would be "essentially" known. If the regression line is known exactly, this approximate method becomes exact.

Only in the situation where the ellipsoid lies totally within the rectangle obtained from the first approximate method is the second approximate procedure desirable. This procedure circumscribes the ellipsoid and is easy to use. In fact, both approximate procedures have the advantage of ease of computation.

VI. Acknowledgement

The author is indebted to the late Professor M. A. Girshick and Professor Rupert Miller for their helpful discussions.

VOL. 3, NO. 1 TECHNOMETRICS FEBRUARY, 1961

The Robustness of Life Testing Procedures Derived from the Exponential Distribution

M. ZELEN* AND MARY C. DANNEMILLER

National Bureau of Standards

Almost all the statistical procedures in current use for evaluating the reliability of components or equipment rest on the assumption that the failure times follow the exponential distribution. However, in practical situations one rarely has enough data to determine whether failure times are actually exponential. This paper studies the behavior of several statistical life testing procedures based on the exponential failure law if the true failure law is the Weibull distribution. It is found that these statistical techniques, which are widely used, are very sensitive to departures from initial assumptions. Applying these techniques to life test data when the exponential failure law is not satisfied may result in substantially increasing the probability of accepting components or equipments having poor mean-time-to-failure.

This paper also develops convenient analytic techniques for approximating (i) the distribution of sums of independent random variables, and (ii) the characteristics of sequential procedures, for non-negative random variables. These techniques are based on a general life distribution which consists of an expansion of generalized Laguerre polynomials. Application of these methods to the Weibull distribution yields excellent results.

1. INTRODUCTION

At the present time there is wide interest in the life testing of components or equipments where the time-to-failure or time-between-failures is assumed to follow the exponential distribution. A recent report [1], often referred to as the AGREE report, contained the findings of two task groups which presented statistical procedures for evaluating the reliability of components or equipments from life tests. A. Lieberman [12] has recently presented a host of sequential life testing procedures similar to those presented in the AGREE report. Much of the statistical theory on which these procedures are based was developed in a series of papers by B. Epstein and M. Sobel [5, 6, 7, 8]. Their papers contain work dealing with several different types of life testing plans.**

The key assumption underlying almost all the statistical procedures associated with life testing is that the failure times follow an exponential distribution with

---

* Also at the University of Maryland.

** The Department of Defense has recently published a handbook of life testing plans based on the exponential distribution. The handbook was released after this manuscript was in galley; the reference is:

Sampling Procedures and Tables for Life and Reliability Testing (Based on Exponential Distribution), Quality Control and Reliability Handbook (Interim) H108, Office of the Assistant Secretary of Defense (Supply and Logistics), 1960.

probability density function and cumulative distribution function

$$f(t) = \frac{1}{\theta} e^{-t/\theta}, \qquad t \geq 0 \tag{1}$$

$$F(t) = \int_0^t f(x)\, dx = 1 - e^{-t/\theta}, \qquad t \geq 0 \tag{2}$$

respectively. The function $1 - F(t)$, usually referred to as the reliability function or failure law, will be denoted by $R(t)$; i.e.,

$$R(t) = \Pr\{\text{failure time} > t\} = e^{-t/\theta}, \qquad t \geq 0 \tag{3}$$

In practical situations one rarely has enough data to verify that the exponential failure law truly characterizes the data. Instead it is hoped that any deviations from exponential theory will not seriously affect the procedures which are derived from it. Statistical procedures which are insensitive to departures from basic assumptions have been termed "robust", cf. Box and Anderson [3].

It is the purpose of this paper to investigate the robustness of four representative acceptance sampling procedures derived from the exponential distribution when, in fact, the failure times follow a Weibull distribution with the same mean life. This is done by constructing the operating characteristic (O.C.) curves for these procedures when the parent distribution of failure times is the Weibull distribution having the reliability function

$$R(t) = \int_t^\infty \frac{p x^{p-1}}{\theta} e^{-x^p/\theta}\, dx = \exp(-t^p/\theta) \tag{4}$$

for values of the shape parameter $p = 3/2$, 2, and 3.

The Weibull distribution belongs to a class of distributions which characterize "wear out" failure, e.g., the longer the item has been used, the greater the probability of failure. It has been successfully applied to studies of the reliability of vacuum tubes, cf. Kao [11]. Further, it is difficult to distinguish between it and the exponential distribution for small sample sizes. To illustrate this, Figure 1 depicts the empirical reliability function, plotted on semi-logarithmic paper for samples of 5, 15, and 25 drawn from the exponential and various Weibull distributions. One would expect the plotted points to approximate a straight line only if the failure times actually follow an exponential distribution. Note that the Weibull and the exponential data are indistinguishable in almost all cases. Further discussion on detecting departures from the exponential distribution is contained in Appendix II.

2. Principal Results and Methodology

Quality of items undergoing life tests is usually designated by specifying the mean-time-between-failures (MTBF) or mean failure time. In choosing an appropriate procedure for evaluating life tests one must take into account (i) the probability of rejecting items which come from a process having the desired MTBF (termed the producer's risk); and (ii) the probability of accepting items coming from a process with a sub-standard MTBF (termed the consumer's risk).

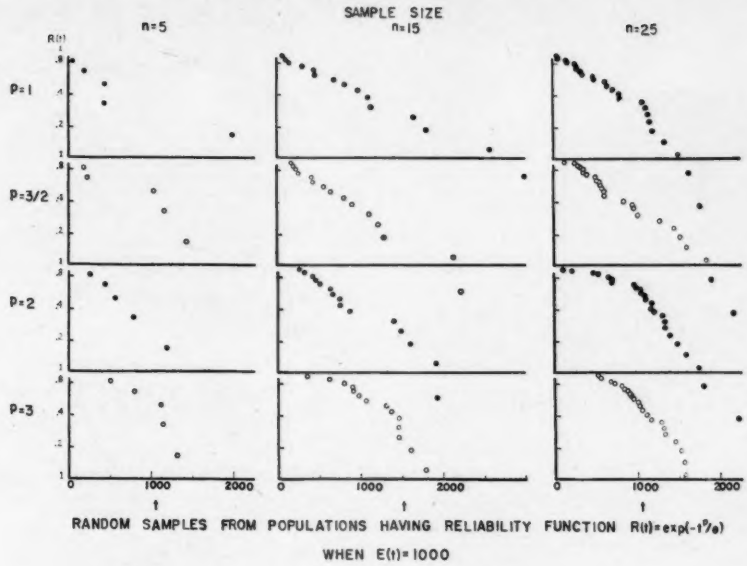


FIGURE 1.

The report of Task Group 2 in AGREE [1] recommended a statistical procedure for life tests which would result in (i) rejecting items having the desired MTBF with probability $\alpha = .10$, and (ii) accepting items having half of the desired MTBF with probability $\beta = .10$. In our investigation we have employed procedures using these risk levels. For purposes of exposition we have taken the desired MTBF to be 1000 hours.

The four life testing plans studied are (1) fixed sample size, (2) fixed sample size with censoring, (3) truncated non-replacement and (4) sequential. These procedures essentially emanate from the works of Epstein and Sobel and are discussed in sections 2.1–2.4.

The O.C. curves for the fixed sample size and the sequential plans were obtained by analytical methods which are described in Appendix I. Aside from their use in robustness studies, these analytic methods give good approximations to (i) the distribution of sums of Weibull random variables and (ii) the O.C. curve and average sample number for sequential tests made on sums of Weibull random variables. The results for the truncated non-replacement plan were calculated theoretically. For the fixed sample size plan with censoring, individual points on the O.C. curves were obtained empirically by simulating the testing procedure on a computer. Each point was based on 1500 simulated life tests. This method of empirical sampling was also used to check some of the O.C. curves obtained by analytic methods.

2.1 *Fixed sample size plan*

Following a well-known procedure, r components are placed on test simultaneously, and the test continues until *all* r components fail. If $t_{1,r} \leq t_{2,r} \leq$

$\cdots \leq t_{r,r}$ denote the failure times ordered in magnitude, then the components are accepted or rejected according to the decision rule:

$$\hat{\theta}_{r,r} = \frac{\sum_{i=1}^{r} t_{i,r}}{r} \begin{matrix} \geq \\ \\ < \end{matrix} C = 1000\,\chi^2_\alpha(2r)^*/2r \begin{matrix} \rightarrow \text{accept} \\ \\ \rightarrow \text{reject} \end{matrix} \tag{5}$$

When $r = 14$, the risk levels $\alpha = .10$ and $\beta = .10$ are obtained for $C = 1000\,\chi^2_{.10}(28)/28 = 676$.

Let the failure times have a Weibull reliability function (4) with mean m, i.e.

$$m = \frac{\theta^{1/p}\Gamma(1/p)}{p}. \tag{6}$$

Then the probability of acceptance, $Pr\{\hat{\theta}_{r,r} > C\}$, can be approximated by

$$\frac{1}{\Gamma(N)}\int_y^\infty e^{-x}x^{N-1}\,dx - \frac{e^{-y}y^N}{N!}\sum_{s=2}^{2q+1}\frac{\alpha_s}{\binom{N+s-1}{s-1}}L^N_{s-1}(y) \tag{7}$$

where $N = rp$, $y = rCp/m$, $L^N_{s-1}(y)$ are Laguerre polynomials defined by (12), and the constants α_s are given by (28) and tabulated in (49). (The derivation of (7) is discussed in Appendix I).

The individual points on the O.C. curves drawn in Figure 2 were calculated using (7) with $q = 2$. A check on the accuracy of this approximation was made by simulating 1500 life tests on each of three samples drawn from populations having mean-time-to-failure $m = 500$, 750, and 1000 hours. These results are summarized in Table 1. The agreement is excellent and well within the sampling error of the simulations.

2.2 *Fixed sample size plan with censoring*

The fixed sample size procedure can be modified by placing n components on test simultaneously and discontinuing the test after the rth ($r < n$) failure. In this case the decision rule is:

$$\hat{\theta}_{r,n} = \frac{\sum_{i=1}^{r} t_{i,n} + (n-r)t_{r,n}}{r} \begin{matrix} \geq \\ \\ < \end{matrix} C = 1000\,\chi^2_\alpha(2r)/2r \begin{matrix} \rightarrow \text{accept} \\ \\ \rightarrow \text{reject} \end{matrix} \tag{8}$$

where $t_{1,n} \leq t_{2,n} \leq \cdots \leq t_{r,n}$ denote the r failure times. This procedure was investigated for $n = 28$, $r = 14$, and $C = 1000\,\chi^2_{.10}(28)/28 = 676$.

Epstein and Sobel show that when the failure times follow the exponential distribution, the distribution of $\hat{\theta}_{r,n}$ is identical with the distribution of $\hat{\theta}_{r,r}$. Hence the operating characteristic curve based on the lowest r out of n ordered observations is identical with that based on r out of r observations.

---

* $\chi^2_\alpha(2r)$ denotes the upper α percentile of the chi-square distribution for $2r$ degrees of freedom.

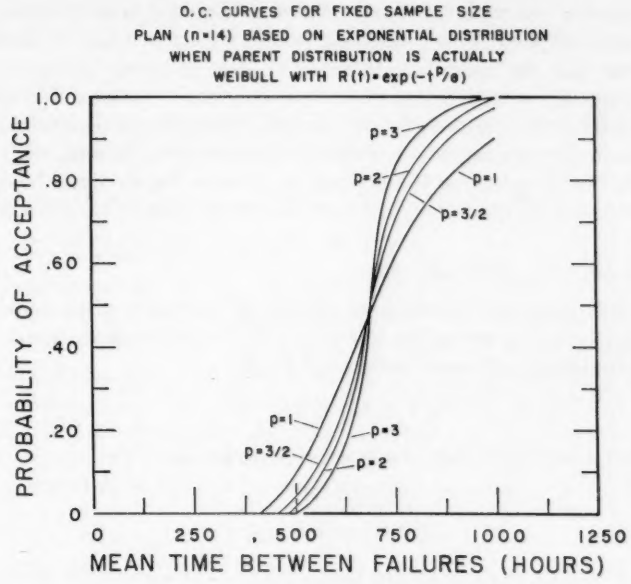


FIGURE 2.

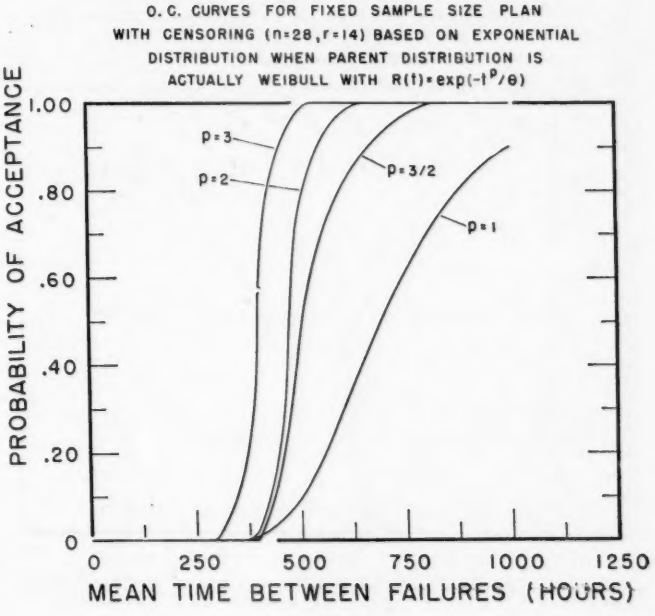


FIGURE 3.

The operating characteristic curve shown in Figure 3 was obtained entirely by empirical sampling. These sampling results are summarized in Table 2.

It is clear that the operating characteristic curves for the censored and uncensored cases are very different when the failure times follow a Weibull distribution instead of the exponential. The censored procedure seems to be particularly sensitive to departures from exponential assumptions. In fact, when $p = 3$, the probability of accepting components with mean failure times between 500 and 1000 hours is so great that such components are almost always accepted.

2.3 *Truncated non-replacement plan*

A truncated non-replacement plan consists of placing n components on test simultaneously and stopping the test after either a preassigned time T or after the rth failure time, whichever occurs sooner. If

$$t_{1,1} \leq t_{2,n} \leq \cdots \leq t_{r,n}$$

denote the failure times, then the decision procedure is:

(i) accept if $t_{r,n} \geq T$
(ii) reject if $t_{r,n} < T$.

Let $R(T)$ denote the probability that a component survives to time T. Then the probability that a sample of size n has at least r failure times larger than T (when the mean failure time is m) is the probability of acceptance; i.e.

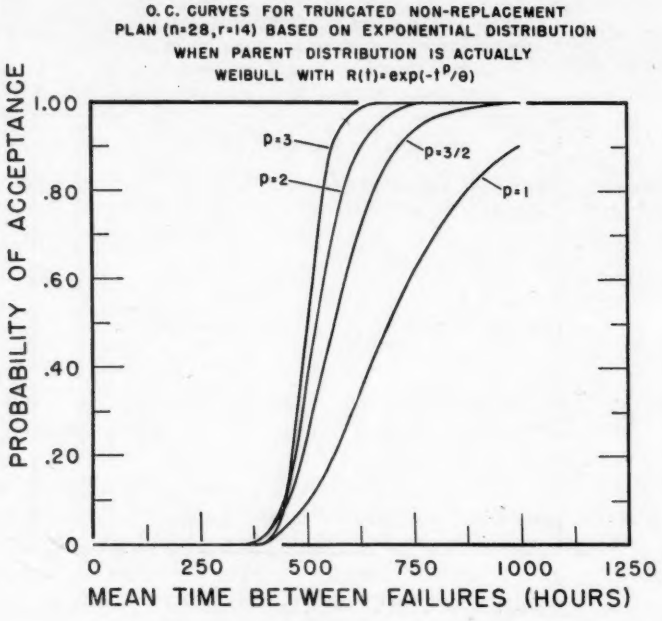


FIGURE 4.

$$L_p(m) = \sum_{s=r}^{n} \binom{n}{s} R(T)^s [1 - R(T)]^{n-s} = I_{R(T)}(r, n - r + 1) \tag{9}$$

where $I_x(a, b)$ is the incomplete beta function.

Assuming the exponential reliability function $R(T) = \exp\{-T/\theta\}$ and taking $n = 28$, $r = 14$, and $T = 504$ hours, we find that

$$L_1(1000) = 0.900$$

$$L_1(500) = 0.095$$

which correspond to the chosen consumer and producer risks. (The consumer's risk is 0.095 rather than 0.100 due to the necessity of r being integral).

Hence it is an easy matter to construct the operating characteristic curve of this procedure for any parent distribution by substituting the appropriate reliability function for different values of m in equation (9). The O.C. curves for Weibull parent distributions using $R(T) = \exp\{-T^p/\theta\}$ are portrayed in Figure 4. Note that like the censored case, the O.C. curves for the Weibull are widely separated from that of the exponential indicating a particularly non-robust procedure.

2.4 *Sequential plan*

The simplest application of a sequential testing procedure consists of testing the components one at a time. After each failure a decision is made either to accept or reject the components, or to run another test.

Let $t_1, t_2, \cdots, t_r, \cdots$ represent the failure times which are not necessarily ordered in magnitude. The sequential procedure used here is a sequential probability ratio test [Wald, 15], [Epstein and Sobel, 8] for testing a simple hypothesis ($H_0 : m = m_0 = 1000$ hours) against a simple alternative ($H_1 : m = m_1 = 500$ hours). The decision rule to be applied after the rth failure is if:

$$\sum_{1}^{r} t_i \geq A(r) \rightarrow \quad \text{accept}$$

$$\sum_{1}^{r} t_i \leq R(r) \rightarrow \quad \text{reject}$$

$$R(r) < \sum_{1}^{r} t_i < A(r) \rightarrow \quad \text{run another life test}$$

where

$$A(r) = 1000\left\{(r - 1)\,\ell n\left(\frac{m_0}{m_1}\right) - \ell n\left(\frac{\beta}{1 - \alpha}\right)\right\}$$

$$R(r) = 1000\left\{r\,\ell n\left(\frac{m_0}{m_1}\right) - \ell n\left(\frac{1 - \beta}{\alpha}\right)\right\}$$

and $\alpha = \beta = .10$.

Figure 5 depicts the O.C. curves for this procedure when the failure times follow the Weibull distribution. These curves were calculated using a method

due to Wald [15]; namely, that the probability of acceptance for a given mean m is

$$L_p(m) \doteq \frac{A^h - 1}{A^h - B^h} \; (h \neq 0), \qquad A = \frac{1-\beta}{\alpha}, \qquad B = \frac{\beta}{1-\alpha} \tag{10}$$

where

$$m = \frac{pw \, \ell n \, (m_0/m_1)}{K(w-1) \ln \left\{ (1-w)^p \sum_{s=0}^{4} a_s \binom{s+p-1}{s} w^s \right\}} \tag{11}$$

$$K = \left(\frac{1}{m_1} - \frac{1}{m_0} \right), \qquad h = \frac{pw}{Km(1-w)}.$$

Equation (10) is derived in Appendix I, and the constants a_s are tabulated in (47). The procedure for calculating $L_p(m)$ is to choose an arbitrary $w < 1$, solve for m, then find h, and finally substitute in (10). To check the accuracy of this approximation, a simulation study for the sequential procedure was carried out on a computer. The sequential tests were truncated at 112 observations. These results are summarized in Table 3.

3. Discussion

None of the four life testing procedures studied in this paper is robust with respect to Weibull alternatives. In particular the censored life test and the truncated non-replacement test are strikingly non-robust. It is obvious from the graphs of the O.C. curves that lots having low mean failure times have a high probability of acceptance when the failure times follow a Weibull distribution with shape parameter $p > 1$. This tendency is increased as p increases. Both these life testing plans depend on the distribution of a function of the order statistics which in turn is sensitive to the original distribution. On the other hand, the fixed sample size plan and the sequential procedure make use of averages. Hence the central limit theorem suggests that these procedures would not be as sensitive to departures from the original distribution. This is borne out by our calculations.

We have tried to show that dogmatic use of life testing procedures without a careful verification of the assumption that failure times follow the exponential distribution may result in a high probability of accepting "poor quality" equipment.

"Quality" has been defined by the mean-time-to-failure or mean-time-between-failures which is simply the first moment of the life distribution. When the failure times follow the exponential distribution, the mean time is a satisfactory criterion for defining quality. However, it may be highly artificial and often misleading for other failure distributions. A more meaningful and practical approach would be to specify a point on the reliability curve. In many instances, the engineer has in mind a percentile, which is transformed into a mean failure time assuming the exponential distribution. For example, one could specify that at least 90% of the items should (on the average) survive to 105 hours,

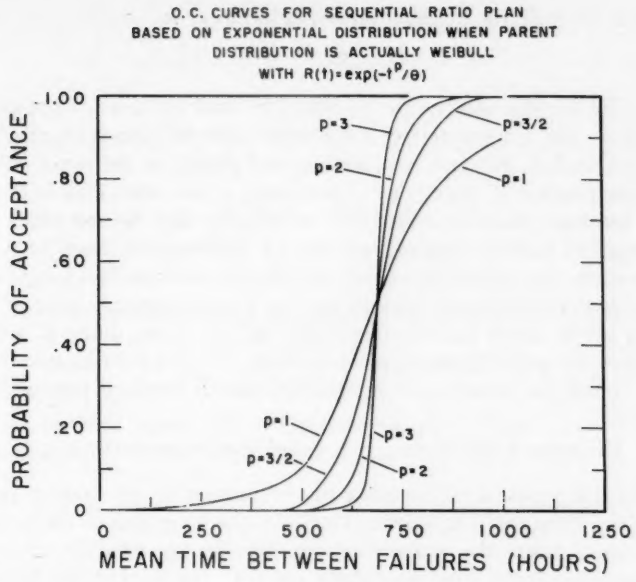


FIGURE 5.

i.e., $R(105) \geq 0.90$. When the reliability function is exponential, this requirement is the same as stating that the mean failure time be at least 1000 hours. However, if the reliability function is Weibull, the requirement on the mean-failure-time will be much different. Note that the mean of the Weibull reliability function

$$R(t) = e^{-t^p/\theta} \quad \text{is} \quad m = \theta^{1/p}\frac{\Gamma(1/p)}{p}.$$

Hence $R(t)$ can be written explicitly as a function of m; i.e.

$$R(t) = \exp\left\{-\left(\frac{t\Gamma(1/p)}{mp}\right)^p\right\}.$$

Thus the minimum value of m for which $R(105) \geq .90$ is

$$m = [-\ln (.90)]^{-1/p}(105)\Gamma(1/p)/p.$$

The corresponding values of the mean failure time for $p = 3/2$, 2, and 3 are 425, 287, and 199 hours respectively. Obviously a low mean failure time does not indicate poor quality in this instance as these values all insure that at least 90% of the items (on the average) will survive to 105 hours.

Alternatively one might specify a different point on the reliability curve. Suppose one specified that at least half the items should survive to 693 hours, i.e. $R(693) \geq 0.50$. If the reliability function is exponential, this will be satisfied for a mean failure time of 1000 hours as before. However, when the reliability

function is Weibull, the minimum mean failure time is

$$m = [-\ell n\ (.50)]^{-1/p}(693)\Gamma(1/p)/p.$$

Values of m for this case are $m = 800$, 738, and 700 hours corresponding to $p = 3/2$, 2, and 3 respectively. Thus when different points on the reliability curve are specified, different requirements are placed on the mean failure time.

The specification of quality by a percentile is not new. This is the practice for ball bearings where the "load life" is actually that life for which 90% (on the average) of ball bearings will survive; c.f. Lieblein and Zelen [13]. Adoption of a percentile for specifying quality should also be done for components and parts. Instead of specifying a mean life one should perhaps speak of "median life" (the life at which half the items survive). Of course different applications may require the specification of different percentiles. In a future paper we hope to develop further the idea of setting specifications in terms of percentiles.

4. Appendix I: Analytical Methods for Studying Robustness

The general problem of studying the robustness of life testing procedures based on the exponential distribution involves the investigation of the operating characteristics when the underlying distribution (or reliability function) is other than exponential. One may study the O.C. curve when the failure times follow a specified distribution; or without choosing particular distributions it is possible to study the robustness of life testing procedures by specifying the values of skewness and excess (γ_1, γ_2).

The primary aim of our robustness studies is to gather qualitative information on the behavior of certain life testing procedures when the distribution of failure times is not exponential. We have found it convenient to develop a general life distribution which can be used both when some of the lower order moments differ from that of the exponential and also when specific alternative distributions are specified for the failure times. In this latter situation, the parameters of the general life distribution are chosen to approximate the specified distribution. This section develops the properties of such a general life distribution which lends itself to easy computation.

4.1 *Laguerre Polynomials*

Our general life distribution makes extensive use of generalized Laguerre polynomials*. The Laguerre polynomial of order α $(\alpha > -1)$ and degree s in x is denoted by $L_s^{\alpha}(x)$. For the convenience of the reader, we here summarize some useful properties of Laguerre polynomials, cf. Erdélyi et al. [9].

$$L_s^{\alpha}(x) = \sum_{j=0}^{s} \binom{s+\alpha}{s-j} \frac{(-x)^j}{j!} \tag{12}$$

$$s!\, L_s^{\alpha}(x) = e^x x^{-\alpha} \frac{d^s}{dx^s}(e^{-x}x^{s+\alpha}) \tag{13}$$

* A referee has called our attention to unpublished work of Gupta and Sobel [10] who also make use of Laguerre polynomials in a life testing context.

$$\int_0^\infty e^{-x}x^\alpha L_r^\alpha(x)L_s^\alpha(x)\,dx = \begin{cases} \dfrac{(\alpha+s)!}{s!} & \text{if } r = s \\ 0 & \text{if } r \neq s \end{cases} \tag{14}$$

$$\int_0^\infty e^{-\tau x}x^\alpha L_s^\alpha(x)\,dx = \frac{(\alpha+s)!\,(\tau-1)^s}{s!\,\tau^{\alpha+s+1}}, \qquad \operatorname{Re}\alpha > -1, \qquad \operatorname{Re}\tau > 0 \tag{15}$$

$$\frac{d}{dx}L_s^\alpha(x) = -L_{s-1}^{\alpha+1}(x) \tag{16}$$

4.2 *A General Life Distribution*

Let X be a non-negative random variable with probability density function

$$p(x) = \frac{e^{-x}x^\alpha}{\Gamma(\alpha+1)}\sum_{s=0}^{k} a_s L_s^\alpha(x) \tag{17}$$

where k is a positive integer and a_s and α are suitably chosen constants with the provision that $a_0 = 1$. The reliability function $R(t)$ can be written

$$R(t) = \int_t^\infty p(x)\,dx = \frac{1}{\Gamma(\alpha+1)}\int_t^\infty e^{-x}x^\alpha\,dx - \frac{e^{-t}t^{\alpha+1}}{\Gamma(\alpha+1)}\sum_{s=1}^{k}\frac{a_s}{s}L_{s-1}^{\alpha+1}(t). \tag{18}$$

Note that when $a_s = 0$, $(s > 0)$ and $\alpha = 0$,

$$p(x) = e^{-x} \quad \text{and} \quad R(t) = e^{-t}$$

which are respectively the p.d.f. and reliability function of the exponential distribution.

If there exists another non-negative random variable having p.d.f. $f(x)$, a suitable approximation for $f(x)$ can be found by choosing a_s such that

$$\int_0^\infty L_s^\alpha(x)f(x)\,dx = \int_0^\infty L_s^\alpha(x)p(x)\,dx = a_s\binom{\alpha+s}{s} \quad \text{for} \quad s = 1, 2, \cdots, k. \tag{19}$$

Assuming the existence of moments to order k for $f(x)$, the definition of a_s by (19) results in the moments of $p(x)$ to order k being equal to those of $f(x)$. The parameter α is chosen such that

$$\frac{f(x)}{p(x)} = O(1) \quad \text{as} \quad x \to 0.$$

This insures that the order of contact for both distributions near the origin is the same.

4.3 *Moments and the Characteristic Function*

Using equation (15) the characteristic function of X is

$$\phi(t) = \int_0^\infty e^{ixt}p(x)\,dx = \sum_{s=0}^{k} a_s\binom{\alpha+s}{s}(-it)^s(1-it)^{-(s+\alpha+1)}. \tag{20}$$

Since we can write

$$(-it)^s(1-it)^{-(s+\alpha+1)} = (-1)^s\sum_{r=\alpha+s}^{\infty}\binom{r}{\alpha+s}(it)^{r-\alpha}$$

the characteristic function can be displayed as

$$\phi(t) = \sum_{r=0}^{\infty} \frac{(it)^r}{r!} \left\{ \frac{(r+\alpha)!}{\alpha!} \sum_{s=0}^{\min(r,k)} (-1)^s a_s \binom{r}{s} \right\}. \tag{21}$$

Thus the moments about the origin of X are

$$E(X^r) = \frac{(r+\alpha)!}{\alpha!} \sum_{s=0}^{\min(r,k)} (-1)^s a_s \binom{r}{s}. \tag{22}$$

In particular the first four moments of X are:

$$\begin{aligned} E(X) &= \frac{(1+\alpha)!}{\alpha!}[1 - a_1] \\ E(X^2) &= \frac{(2+\alpha)!}{\alpha!}[1 - 2a_1 + a_2] \\ E(X^3) &= \frac{(3+\alpha)!}{\alpha!}[1 - 3a_1 + 3a_2 - a_3] \\ E(X^4) &= \frac{(4+\alpha)!}{\alpha!}[1 - 4a_1 + 6a_2 - 4a_3 + a_4]. \end{aligned} \tag{23}$$

Hence if μ_r' denotes the rth moment about the origin of a known random variable W, one could approximate the distribution of W by letting the moments of W equal the moments of X. Then the constants a_s ($s \leq 4$) are:

$$\begin{aligned} a_1 &= 1 - m_1 \\ a_2 &= \frac{(m_2 - 2) + 4(1 - m_1)}{2} \\ a_3 &= \frac{(6 - m_3) + 9(m_2 - 2) + 18(1 - m_1)}{6} \\ a_4 &= \frac{(m_4 - 24) + 16(6 - m_3) + 72(m_2 - 2) + 96(1 - m_1)}{24} \end{aligned} \tag{24}$$

where

$$m_s = \frac{\mu_s'}{\binom{s+\alpha}{s}}.$$

For the generalized p.d.f. $p(x)$, it is sometimes convenient to set $a_1 = 0$ and consider the distribution of the random variable

$$Y = \left(\frac{m}{1+\alpha}\right) X, \qquad (m > 0).$$

If $a_1 = 0$, $E(X) = (1 + \alpha)$ and $E(Y) = m$. Then the p.d.f. of Y is

$$p(y) = \left(\frac{1+\alpha}{m}\right) p(x) = \left(\frac{1+\alpha}{m}\right) \frac{e^{-x}}{\Gamma(\alpha+1)} x^{\alpha} \sum_{s=0}^{k} a_s L_s^{\alpha}(x), \tag{25}$$

where

$$x = \left(\frac{1+\alpha}{m}\right)y$$

with characteristic function

$$\phi_s(t) = \sum_{s=0}^{k} a_s \binom{\alpha+s}{s}\left(-\frac{im}{1+\alpha}\,t\right)^s\left(1-\frac{im}{1+\alpha}\,t\right)^{-(s+\alpha+1)}. \tag{26}$$

The constants a_s can be found from (24) with $m_1 = 1$.

4.4 *Distribution of a Sum*

Let $X_1, X_2, \cdots, X_n$ be n independent random variables having the p.d.f. $p(x)$. Then the characteristic function of the sum $\xi = \sum_{i=1}^{n} X_i$ is given by

$$\phi_\xi(t) = [\phi_x(t)]^n = (1-it)^{-n(1+\alpha)}\left[\sum_{s=0}^{k} a_s\binom{\alpha+s}{s}(-it)^s(1-it)^{-s}\right]^n$$

which after expansion and some simplification can be written

$$\phi_\xi(t) = (1-it)^{-n(1+\alpha)}\sum_{s=0}^{kn}\alpha_s(-it)^s(1-it)^{-s} \tag{27}$$

where α_s is the coefficient of z^s in

$$\left\{1+\binom{\alpha+1}{1}a_1z+\cdots+\binom{\alpha+k}{k}a_kz^k\right\}^n;\quad \text{explicitly}$$

$$\alpha_s = \sum_{s_1+2s_2+\cdots+ks_k=s}\frac{n!}{s_1!\,s_2!\cdots s_k!\,(n-s_1-\cdots-s_k)!}\cdot\left[\binom{\alpha+1}{1}a_1\right]^{s_1}\cdots\left[\binom{\alpha+k}{k}a_k\right]^{s_k}. \tag{28}$$

The p.d.f. of ξ, say $g(\xi)$, is found by using the inversion formula for characteristic functions, i.e.

$$g(\xi) = \frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-it\xi}\phi_\xi(t)\,dt = \sum_{s=0}^{kn}\alpha_s\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-it\xi}(-it)^s(1-it)^{-(n+n\alpha+s)}\,dt. \tag{29}$$

Since the characteristic function associated with the gamma distribution with p.d.f. $(e^{-x}x^{N+s-1})/\Gamma(N+s)$ is $(1-it)^{-(N+s)}$, we can write

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-itx}(1-it)^{-(N+s)}\,dt = \frac{e^{-x}x^{N+s-1}}{\Gamma(N+s)}.$$

Differentiating the above expression s times with respect to x results in

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-itx}(-it)^s(1-it)^{-(N+s)}\,dt = \frac{1}{\Gamma(N+s)}\frac{d^s}{dx^s}(e^{-x}x^{N+s-1})$$

$$= \frac{s!}{\Gamma(N+s)}L_s^{N-1}(x)e^{-x}x^{N-1}. \tag{30}$$

Letting $N = (n + n\alpha)$ and substituting (30) in (29) gives

$$g(\xi) = \sum_{s=0}^{kn} \frac{\alpha_s s!}{\Gamma(N+s)} L_s^{N-1}(\xi) e^{-\xi} \xi^{N-1}. \tag{31}$$

In order to find the reliability function of ξ, we observe that

$$\frac{s!}{\Gamma(N+s)} \int_t^\infty L_s^{N-1}(x) e^{-x} x^{N-1}\, dx = -\frac{d^{s-1}}{dt^{s-1}} \frac{(e^{-t} t^{N+s-1})}{\Gamma(N+s)} = -\frac{\Gamma(s)}{\Gamma(N+s)} L_{s-1}^N(t) e^{-t} t^N.$$

Hence the reliability function of ξ is

$$R_\xi(t) = \int_t^\infty g(\xi)\, d\xi = \frac{1}{\Gamma(N)} \int_t^\infty e^{-x} x^{N-1}\, dx - \frac{e^{-t} t^N}{N!} \sum_{s=1}^{kn} \frac{\alpha_s L_{s-1}^N(t)}{\binom{N+s-1}{s-1}} \tag{32}$$

where $N = n(1 + \alpha)$.

Now $\alpha_s = O(N^s)$ and from Szego [14, p. 175, eq. 7.68],

$$L_{s-1}^N(x) = x^{-N/2-1/4} O\big((s-1)^{-N/2-1/4}\big);$$

therefore,

$$\alpha_s \frac{\Gamma(s)}{\Gamma(N+s)} L_{s-1}^N(x) e^{-x} x^N = X^{(2N-1)/4} e^{-x} O\big((s-1)^{-N/2-1/4}\big)$$

which indicates that the leading terms of (32) may provide a good approximation to $R_\xi(t)$.

When the random variable under consideration is $Y = Xm/(1 + \alpha)$, the reliability function of $\eta = \sum_{i=1}^n Y_i$ is $R_\eta(t) = R_\xi t(1 + \alpha/m)$. The α_s are defined by (28) with $a_1 = 0$. In this case,

$$\alpha_s = \begin{cases} O(N^{s/2}) & \text{if } s \text{ even} \\ O(N^{(s-1)/2}) & \text{if } s \text{ odd.} \end{cases}$$

Hence for small s

$$\alpha_s \frac{\Gamma(s)}{\Gamma(N+s)} L_{s-1}^N(x) e^{-x} x^N = \begin{cases} O(N^{-s/2}) & \text{if } s \text{ even} \\ O(N^{-(s-1)/2}) & \text{if } s \text{ odd} \end{cases}$$

and

$$R_\eta(t) = \frac{1}{\Gamma(N)} \int_u^\infty e^{-x} x^{N-1}\, dx - \frac{e^{-u} u^N}{N!} \sum_{s=2}^{2q+1} \frac{\alpha_s L_{s-1}^N(\mu)}{\binom{N+s-1}{s-1}} + O(N^{-(q+1)}) \tag{33}$$

where $u = t(1 + \alpha)/m$ and $N = n(1 + \alpha)$.

4.5 *Sequential Tests*

B. Epstein and M. Sobel [8] have developed several sequential life test procedures based on the exponential distribution. Their development uses Wald's work [15] virtually without modification. In this section we develop the necessary

tools for studying the behavior of one of these procedures when the underlying probability density function is the general life p.d.f. given by (17).

Let $X_1, X_2, \cdots, X_r, \cdots$ denote observations of failure times from the exponential distribution

$$f(x) = \frac{1}{m} e^{-x/m} \qquad x \geq 0.$$

Then the sequential ratio procedure for testing the simple hypothesis H_0 : $m = m_0$ against the simple alternative $H_1 : m = m_1$ $(m_1 < m_0)$ with type I and type II errors equal to α and β is if

$$\begin{aligned} &\ell n\, B < \sum_{i=1}^{r} Z_i < \ell n\, A \Rightarrow \text{take another observation} \\ &\sum_{i=1}^{r} Z_i > \ell n\, A \Rightarrow \text{accept } H_1 \\ &\sum_{i=1}^{r} Z_i < \ell n\, B \Rightarrow \text{accept } H_0 \end{aligned} \tag{34}$$

where

$$\begin{aligned} Z_i &= \ell n \left(\frac{m_0}{m_1}\right) - \left(\frac{1}{m_1} - \frac{1}{m_0}\right) X_i \\ A &= \frac{1-\beta}{\alpha}, \qquad B = \frac{\beta}{1-\alpha}. \end{aligned} \tag{35}$$

Let $L(m)$ be the probability of acceptance when the mean is m. Then Wald [15] has shown that a good approximation to the O.C. curve is given by

$$L(m) \doteq \frac{A^h - 1}{A^h - B^h} \tag{36}$$

where h is the non-zero solution of

$$E_m(e^{Zh}) = 1 \tag{37}$$

and $E_m(\cdot)$ indicates taking the expectation when the mean is m.

Similarly the average (expected) sample number (ASN) to reach a decision can be calculated approximately by

$$E_m(n) \doteq \frac{L(m)\, \ell n\, B + [1 - L(m)]\, \ell n\, A}{E_m(Z)}. \tag{38}$$

Both (36) and (38) are only valid when $E_m(Z) \neq 0$.

Our objective in this section is to evaluate the O.C. curve and ASN when the sequential procedure (34) is used but the underlying p.d.f. of failure times is not exponential but is given by (17).

Now

$$E_m(e^{Zt}) = \left(\frac{m_0}{m_1}\right)^t E_m(e^{-KtX}) = \left(\frac{m_0}{m_1}\right)^t \phi(iKt), \qquad K = \left(\frac{1}{m_1} - \frac{1}{m_0}\right)$$

and

$$\phi(iKt) = E_m(e^{-KtX}) = \sum_{s=0}^{k} a_s \int_0^\infty \frac{x^\alpha L_s^\alpha(x)}{\Gamma(\alpha+1)} e^{-x[1+Kt]}\, dx.$$

The above integral exists provided $1 + Kt > 0$.
Therefore using (37)

$$E(e^{Zh}) = \left(\frac{m_0}{m_1}\right)^h (1+Kh)^{-(1+\alpha)} \sum_{s=0}^{k} a_s \binom{s+\alpha}{s} (Kh)^s (1+Kh)^{-s} = 1. \qquad (39)$$

Since $a_s (s > 0)$ are functions of m [cf. (24)], one substitutes an arbitrary $h \neq 0$ in (39) and finds the corresponding value of m; $L(m)$ and $E_m(h)$ can then be calculated.

For the case where $a_1 = 0$ and

$$Y = \left(\frac{m}{1+\alpha}\right)X, \qquad Z_i = \left[\ell n\left(\frac{m_0}{m_1}\right) - \frac{Km}{1+\alpha} X_i\right]$$

and thus

$$E(e^{Zh}) = \left(\frac{m_0}{m_1}\right)^h \left(1 + \frac{Kmh}{1+\alpha}\right)^{-(1+\alpha)} \cdot \sum_{s=0}^{k} a_s \binom{s+\alpha}{s} \left(\frac{Kmh}{1+\alpha}\right)^s \left(1 + \frac{Kmh}{1+\alpha}\right)^{-s} = 1. \qquad (40)$$

The above expression exists only when $1 + Kmh/(1+\alpha) > 0$.

In order to calculate $L(m)$ and $E_m(n)$ we let

$$w = \left(\frac{Kmh}{1+\alpha}\right)\left(1 + \frac{Kmh}{1+\alpha}\right)^{-1}.$$

Then

$$h = \frac{1+\alpha}{Km} \cdot \frac{w}{1-w}$$

and (40) can be written

$$\left(\frac{m_0}{m_1}\right)^{[(1+\alpha)/Km][w/(1-w)]} (1-w)^{1+\alpha} \sum_{s=0}^{k} a_s \binom{s+\alpha}{s} w^s = 1. \qquad (41)$$

Taking logarithms and re-arranging terms results in

$$m = \frac{(1+\alpha)w}{K(w-1)} \cdot \frac{\ell n\left(\frac{m_0}{m_1}\right)}{\ell n\left\{(1-w)^{1+\alpha} \sum_{s=0}^{k} a_s \binom{s+\alpha}{s} w^s\right\}}. \qquad (42)$$

Thus one need only choose an arbitrary w, solve for m and using $h = [(1+\alpha)/Km][w/(1-w)]$ find the corresponding values of $L(m)$ and $E_m(n)$. (Since $1 + Kmh/(1+\alpha) > 0$, this requires that $w < 1$).

The case $h = 0$ occurs when $E(Z_i) = \ell n\,(m_0/m_1) - Km = 0$; i.e.

$$m = K^{-1}\, \ell n\,(m_0/m_1).$$

One can then apply the results of Wald [15, p. 176] to get

$$L(m) \doteq \frac{\ell n\, A}{\ell n\, A - \ell n\, B} \tag{43}$$

$$E_m(n) \doteq -\frac{(1+\alpha)\,\ell n\, A\, \ell n\, B}{\left[\ell n \frac{m_0}{m_1}\right]^2 [1 + 2a_2 + 2\alpha_2]}. \tag{44}$$

4.6 *Application to the Weibull Distribution*

This section will apply the various formulas developed previously to approximate the distribution of a sum of independent random variables drawn from the Weibull distribution with p.d.f.

$$f(y) = \frac{p y^{p-1}}{\theta} \exp\{-(y^p/\theta)\}, \qquad y \geq 0.$$

The moments about the origin for the Weibull distribution are easily shown to be

$$E(Y^s) = \mu_s' = \theta^{s/p}\Gamma\left(1 + \frac{s}{p}\right). \tag{45}$$

Since the mean is

$$m = \mu_1' = \theta^{1/p}\Gamma\left(1 + \frac{1}{p}\right)$$

the moments can be expressed

$$\mu_s' = m^s g(s), \qquad g(s) = \frac{\Gamma\left(1 + \frac{s}{p}\right)}{\left[\Gamma\left(1 + \frac{1}{p}\right)\right]^s}. \tag{46}$$

If the order of contact at zero for the generalized p.d.f. $p(x)$ given in (17) is to be the same as for $f(y)$, then we must choose $\alpha = p - 1$. Since the moments μ_s' are powers of m, it is convenient to set $a_1 = 0$ in (17) and consider the random variable $Y = Xm/(1 + \alpha)$ where $E(X) = 1 + \alpha = p$. Then the reliability function of $\sum_i Y_i$ can be approximated by (33) where a_s can be found using expressions like (24) with

$$m_s = \frac{p^s g(s)}{\binom{s + p - 1}{s}}.$$

In approximating $p(x)$ we chose $k = 4$. The values of a_s $(s = 2, 3, 4)$ are tabulated below for $p = 3/2, 2, 3$.

	$p=3/2$	2	3
a_2	−0.12340	−0.15117	−0.15093
a_3	−0.06925	−0.09014	−0.08475
a_4	−0.01879	−0.02122	−0.01211

(47)

The corresponding values of α_s to be used in (33) are given by (28). The first five values are

$$\begin{aligned} \alpha_0 &= 1, & \alpha_1 &= 0 \\ \alpha_2 &= n\binom{p+1}{2}a_2, & \alpha_3 &= n\binom{p+2}{3}a_3 \\ \alpha_4 &= \binom{n}{2}\left[\binom{p+1}{2}a_2\right]^2 + n\binom{p+3}{4}a_4, \\ \alpha_5 &= \frac{n-1}{n}\alpha_2\alpha_3 . \end{aligned} \tag{48}$$

This results in the numerical values

	$p=3/2$	2	3
α_2	−3.23927	−6.34929	−12.67812
α_3	−2.12089	−5.04788	−11.86468
α_4	−0.64042	17.23179	49.20474
α_5	+6.37942	29.76115	139.67745

(49)

The constants tabulated in (48) and (49) were used to calculate the O.C. curves for the fixed sample size and sequential procedures described in section 2.

5. Appendix II: Tests for the Validity of the Exponential Assumption

B. Epstein in a recent paper [4] enumerates several procedures for testing the validity of the assumption that observations come from an exponential distribution. In connection with our robustness study, a brief investigation was undertaken to see how powerful these procedures are in distinguishing between the exponential and Weibull distributions. The two tests referred to by Epstein as "Test 3" and "Test 8" were examined. Test 3 is conditional on the total time on test; Test 8 was chosen because Bartholomew [2] compared the power of three competing tests and found Test 8 to be the asymptotically most powerful of the three against Weibull alternatives.

Suppose we are interested in testing the hypothesis that the failure times $t_1 \leq t_2 \leq \cdots \leq t_r$ come from an exponential distribution. Consider a non-replacement censored situation where the life test starts with n items and ends as soon as the first r failures occur ($r \leq n$). Then the total time on test up to and including the ith failure is

$$T(t_i) = t_1 + t_2 + \cdots + t_{i-1} + (n - i + 1)t_i , \; i = 1, \cdots , r$$

i.e.

$$T(t_1) = nt_1 , \qquad T(t_2) = t_1 + (n-1)t_2 , \cdots .$$

Then the conditional distribution of $T(t_1)$, $T(t_2)$, $\cdots$, $T(t_{r-1})$ is uniform over the interval $[0, T(t_r)]$ for fixed $T(t_r)$. Hence

$$Z = \frac{\sum_{i=1}^{r-1} T(t_i) - \frac{(r-1)}{2} T(t_r)}{T(t_r)\sqrt{\frac{r-1}{12}}}$$

is an approximate standard normal deviate. Epstein (Test 3) suggests that we reject at level α if $|Z| > Z_\alpha$ where α is the $100(1 - \alpha/2)$ percentage point of the standard normal distribution.

Epstein's Test 8 uses the statistic

$$W = 2r \frac{\left\{ \ell n \frac{T(t_r)}{r} - \frac{1}{r} \ell n \left[T(t_1) + \ell n\, T(t_2 - t_1) + \cdots + \ell n\, T(t_r - t_{r-1}) \right] \right\}}{1 + \frac{r+1}{6r}}$$

where $T(t_i - t_{i-1})$ is the total time on test in the interval $[t_{i-1}, t_i]$. This statistic is asymptotically distributed as a χ^2 variable with $(r - 1)$ degrees of freedom. Hence the test procedure is to reject at level α if $W > \chi^2_\alpha(r - 1)$.

In order to determine the operating characteristics of these tests, i.e., the proportion of times the tests would distinguish between the exponential and Weibull distributions, an empirical sampling study was run on an electronic computer. An $\alpha = .05$ level of significance was chosen. Random samples were drawn from each of the Weibull distributions with shape parameter $p = 3/2$, 2, and 3 and also from the exponential distribution. All distributions had the first moment equal to unity. Both testing procedures were simulated on samples from these distributions; the results of this simulation are tabulated below.

Proportion of times exponential assumption rejected at .05 level of significance

Test	p	$n = 14, r = 14$	$n = 25, r = 14$	$n = 25, r = 25$
3	1	.057	.050	
	3/2	.340	.253	
	2	.837	.643	
	3	1.000	.983	
8	1	.043*	.063	.045*
	3/2	.063*	.080	.120*
	2	.178*	.137	.325*
	3	.575*	.420	.847*

* Starred values are based on 600 samples or simulations. All other values are based on 300 samples.

It is obvious for small sample sizes that Test 8 is not powerful enough to distinguish between the two distributions. Interestingly Test 3, which is conditional upon the total time on test, appears to be more powerful than Test 8, but even it is effective only when the shape parameter p becomes large.

TABLE 1
Probability of Acceptance for Fixed Sample Size Plan
where $R(t) = e^{-tp/\theta}(r = 14)$

MTBF (Hours)	Empirical Sampling Results* p				Analytical Results p		
	1	3/2	2	3	3/2	2	3
1000	.906	.973	.996	1.000	.975	.995	1.002
750	.615	.705	.743	.837	.691	.749	.830
500	.103	.039	.010	.000	.037	.014	−.003

TABLE 2
Probability of Acceptance for Censored Plan
where $R(t) = e^{-tp/\theta}(n = 28, r = 14)$

MTBF (Hours)	Empirical Sampling Results* p			
	1	3/2	2	3
1000	.893	1.000	1.000	1.000
750	.595	.971	1.000	1.000
600	—	.814	.977	1.000
500	.106	.463	.777	.985
400	—	.079	.190	.516
300	—	.000	.000	.000

TABLE 3
Probability of Acceptance for a Sequential Plan
where $R(t) = e^{-tp/\theta}$ (Truncated at 112 trials)

MTBF (Hours)	Empirical Sampling Results* p			
	1	3/2	2	3
1000	.927	.995	.998	1.000
750	.645	.777	.884	.973
690	.508	.524	.554	.542
660	.437	.374	.319	.157
600	.287	.147	.048	.001
500	.097	.013	.000	.000

* Each entry is based on 1500 simulated life tests.

References

[1] Advisory Group on Reliability of Electronic Equipment. *Reliability of Military Electronic Equipment*, Office of the Assistant Secretary of Defense (Research and Engineering), 1957.

[2] Bartholomew, D. J. "Testing for departure from the exponential distribution", *Biometrika, 44*, (1957), 253–256.

[3] Box, G. E. P. and Anderson, S. L. "Permutation theory in the derivation of robust criteria and the study of departures from assumption", *Journal of the Royal Statistical Society, Series B, 17* (1955), 1–26.

[4] Epstein, B. "Testing for the validity of the assumption that the underlying distribution of life is exponential, Part I", *Technometrics, 2* (1960), 83–101.

[5] Epstein, B. "Truncated life tests in the exponential case", *The Annals of Mathematical Statistics, 25* (1954), 555–564.

[6] Epstein, B. and Sobel, M. "Life testing", *Journal of the American Statistical Association, 48* (1953), 486–502.

[7] Epstein, B. and Sobel, M. "Some theorems relevant to life testing from an exponential distribution", *The Annals of Mathematical Statistics, 25* (1954), 373–381.

[8] Epstein, B. and Sobel, M. "Sequential life tests in the exponential case," *The Annals of Mathematical Statistics, 26* (1955), 82–93.

[9] Erdélyi, A., Magnus, W., Oberhettinger, F. and Tricomi, F., *Higher Transcendental Functions*, Volume 2, New York: McGraw-Hill Book Co., Inc., 1953.

[10] Gupta, S. S. and Sobel, M. "Laguerre expansions in life testing", unpublished manuscript.

[11] Kao, J. H. K. "Computer methods for estimating Weibull parameters in reliability studies", *IRE Transactions on Reliability and Quality Control*, PGRQC-13 (1958), 15–22.

[12] Lieberman, A. "Sequential life test plans for the exponential distribution", *Industrial Quality Control, XVI* (1959), 14–18.

[13] Lieblein, J. and Zelen, M. "Statistical investigation of the fatigue life of deep-groove ball bearings", *Journal of Research of the National Bureau of Standards, 57* (1956), 273–316.

[14] Szego, G. "Orthogonal polynomials", *American Mathematical Society Colloquim Publications, XXIII* (1959).

[15] Wald, A. *Sequential Analysis*, New York: John Wiley and Sons, Inc., 1947.

[16] Zelen, M. and Dannemiller, Mary C. "Are life testing procedures robust?" *Proceedings of the Sixth National Symposium on Reliability and Quality Control* (1960), 185–189.

VOL. 3, NO. 1 TECHNOMETRICS FEBRUARY, 1961

An Application of a Balanced Incomplete Block Design

PETER W. M. JOHN

California Research Corporation, Richmond, California
and the University of California, Berkeley

The analysis of balanced incomplete block experiments is discussed in most of the standard textbooks on experimental design. These discussions are usually confined to qualitative treatments; it being customary to obtain an adjusted sum of squares for treatments and to give procedures for determining the significance of the observed difference between two treatment totals.

This paper describes a balanced incomplete block experiment in which the nine treatments were quantitative rather than qualitative, being actually two additives each at four levels and a third at one level. The unusual feature of the analysis is found in Section 3 where the adjusted sum of squares for treatments is subdivided into individual degrees of freedom, each of which is meaningful and specific to this example, and with which we obtain from the data response curves for the two factors which were used at four levels each.

1. INTRODUCTION

In this dishwashing test solutions of the treatments (detergent formulations) are made up, and plates soiled with a standard soil are washed one at a time until they are clean. The testing procedure stipulates that three basins shall be used (three treatments tested at the same time). The three operators wash at the same speed during the test, and the "yield" reported is the number of plates washed before the foam disappears. If we consider each set of three such concurrent trials as a block, it is clear that variation between the blocks can be expected, due to such factors as changes in the temperature of the washing room and in the common washing rate from block to block.

We are specifically concerned with a balanced incomplete block experiment with nine treatments in 12 blocks of three plots each; this design is also a balanced lattice.

Three basic detergents were used. Two of them were tested with 0, 1, 2, and 3 parts of an additive. The third was a standard detergent, which was expected to give a higher yield than the other eight treatments.

We shall denote the treatments by letters as follows: A, B, C, D for base Detergent I with 3, 2, 1, 0 parts of additive, respectively, E, F, G, H for base Detergent II with 3, 2, 1, 0 parts of additive, respectively, and J for the control treatment. The design used was (the observed yields are given in parentheses): The 12 blocks were actually run in random order, and the allocation of treatments to basins within blocks was random, so that the experiment was not divisible into replications for the purposes of analysis.

1	A (19)	B (17)	C (11)	7	A (20)	E (26)	J (31)
2	D (6)	E (26)	F (23)	8	B (16)	F (23)	G (21)
3	G (21)	H (19)	J (28)	9	C (13)	D (7)	H (20)
4	A (20)	D (7)	G (20)	10	A (20)	F (24)	H (19)
5	B (17)	E (26)	H (19)	11	B (17)	D (6)	J (29)
6	C (15)	F (23)	J (31)	12	C (14)	E (24)	G (21)

2. The Intrablock Analysis

We give in this section a summary of the usual intrablock analysis. For details the reader is referred to such standard works as Davies [1], Scheffé [2].

The model assumes that the yield for the ith treatment in the jth block is given by

$$Y_{ij} = \mu + t_i + \beta_j + e_{ij} \tag{2.1}$$

where μ is the grand mean, t_i the ith treatment effect, β_j the jth block effect, and e_{ij} the intrablock error with $\sum_i t_i = \sum_j \beta_j = 0$. The errors e_{ij} are independently normally distributed with zero mean and variance σ^2. There are k experimental units in each block and a total of t treatments. Each treatment is repeated r times and each pair of treatments occurs in the same block λ times, where $(t-1)\lambda = r(k-1)$.

We compute for each treatment quantities Q_i, where $Q_i = k\cdot$(total of all units with the ith treatment)—(total of all blocks containing the ith treatment). Then $\hat{t}_i = Q_i/\lambda t$, var $(Q_i) = rk(k-1)\sigma^2$ and cov $(Q_i, Q_{i'}) = -\lambda k\sigma^2$. The adjusted sum of squares between treatments is $\sum_i Q_i^2/\lambda kt$. For this experiment where $k = 3$, $t = 9$, $r = 4$, $\lambda = 1$, we have $9\hat{t}_i = Q_i$, var $Q_i = 24\sigma^2$, cov $(Q_i, Q_{i'}) = -3\sigma^2$, $i, i' = A, B, \cdots, J$, while $Q_A = 3$, $Q_B = -20$, $Q_C = -56$, $Q_D = -116$, $Q_E = 53$, $Q_F = 32$, $Q_G = 15$, $Q_H = -2$ and $Q_J = 91$. The mean square for error is 0.82 (16 d.f.).

3. Responses of Detergents to the Additive

We wish now to go further than the analysis usually given in the standard texts for qualitative variables and to consider the nature of the responses of the two base detergents to the additive. The adjusted sum of squares between treatments will be subdivided into eight components, each having one degree of freedom corresponding to a contrast between treatments.

We rewrite the model letting the t_i of (2.1) be expressed as:

$$t_i = a_h + b_h u_1(x_i) + c_h u_2(x_i) + d_h u_3(x_i) \tag{3.1}$$

where $h = 1$ for $i = A, B, C, D$, $h = 2$ for $i = E, F, G, H$, x_i is the amount of additive ($x_i = 3, 2, 1, 0$) in the ith treatment and u_1, u_2, u_3 are orthogonal polynomials in x_i. Thus, $u_1(x) = 2x - 3 = +3, +1, -1, -3$; $u_2 = (u_1^2 - 5)/4 = +1, -1, -1, +1$; $u_3 = (5u_1^3 - 41u_1)/12 = +1, -3, +3, -1$. Then $\mu + a_h$ represents the expected value of the average yield for the hth base detergent over the four levels of the additive.

For the control we write $t_j = a_3$ and have

$$Y_{Jj} = \mu + a_3 + \beta_j + e_{Jj} \tag{3.2}$$

We choose eight orthogonal contrasts between the Q_i, each corresponding to one of the eight degrees of freedom between treatments.[1] The contrasts are given in Table I, together with their expectations, variances, and observed

TABLE I
Contrasts

	Contrast	Exp. Value	Variance	Obs. Value
1. Linear Component I	$3Q_A + Q_B - Q_C - 3Q_D$	$180\, b_1$	$540\sigma^2$	393
2. Quadratic Component I	$Q_A - Q_B - Q_C + Q_D$	$36\, c_1$	$108\sigma^2$	−37
3. Cubic Component I	$-Q_A + 3Q_B - 3Q_C + Q_D$	$180\, d_1$	$540\sigma^2$	−11
4. Linear Component II	$3Q_E + Q_F - Q_G - 3Q_H$	$180\, b_2$	$540\sigma^2$	182
5. Quadratic Component II	$Q_E - Q_F - Q_G + Q_H$	$36\, c_2$	$108\sigma^2$	4
6. Cubic Component II	$-Q_E + 3Q_F - 3Q_G + Q_H$	$180\, d_2$	$540\sigma^2$	−4
7. I vs II	$\sum_A^D Q - \sum_E^H Q$	$36\,(a_1 - a_2)$	$216\sigma^2$	−287
8. Control vs I, II	$Q_J = -\sum_A^H Q$	$4\,(2a_3 - a_1 - a_2)$	$24\sigma^2$	91

values. They correspond to linear, quadratic, and cubic components for the responses of base Detergents I and II, a comparison between I and II, and a comparison between the control and the average of I and II.

We may easily show by a simple substitution of the appropriate x_i values in (3.1) that

$$E(Q_A) = 9a_1 + 27b_1 + 9c_1 + 9d_1 \tag{3.3}$$

$$E(Q_B) = 9a_1 + 9b_1 - 9c_1 - 27d_1 \tag{3.4}$$

$$E(Q_C) = 9a_1 - 9b_1 - 9c_1 + 27d_1 \tag{3.5}$$

$$E(Q_D) = 9a_1 - 27b_1 + 9c_1 - 9d_1 \tag{3.6}$$

It follows that $E\,(3Q_A + Q_B - Q_C - 3Q_D) = 180b_1$. Furthermore since var $Q_i = 24\sigma^2$ and cov $(Q_i, Q_j) = -3\sigma^2$, var $(-3Q_A - Q_B + Q_C + 3Q_D) = 540\sigma^2$. The details are similar for the other contrasts.

The expectation of the last contrast is $E(Q_J) = 9a_3$, but since

[1] Since all the Q_i have the same variance and all pairs $(Q_i, Q_{i'})$ $(i \neq i')$ the same covariance orthogonal contrasts in the Q_i are orthogonal contrasts in the observations, and lead to a subdivision of the treatment sum of squares.

$$4a_1 + 4a_2 + a_4 = 0, \qquad E(Q_J) = 9a_3 - (4a_1 + 4a_2 + a_3)$$
$$= 4(2a_3 - a_1 - a_2). \qquad (3.7)$$

The sum of squares for a contrast is obtained by squaring the value of the contrast and dividing by the coefficient of σ^2 in the expression for the variance of the contrast.

This leads to the analysis of variance table.

TABLE II
Analysis of Variance

Source of Variation	Sum of Squares	d.f.	m.s.	F
Linear I	286.02	1	286.02	348.80**
Quadratic I	12.68	1	12.68	15.46**
Cubic I	0.22	1	0.22	0.27
Linear II	61.34	1	61.34	74.80**
Quadratic II	0.15	1	0.15	0.18
Cubic II	0.03	1	0.03	0.04
I vs II	381.34	1	381.34	465.05**
Control vs I, II	345.04	1	345.04	420.78**
Treatments (adj)	1086.81	8		
Blocks (unadj)	412.75	11		
Intrablock Error	13.19	16	0.82	
Total	1512.75	35		

It is clear from this table that the response of base Detergent I to the additive may be represented by a quadratic curve while the response of Base II is linear. The linear relationship for base Detergent II is estimated by

$$Y = \hat{\mu} + \hat{a}_2 + \hat{b}_2(2x - 3)$$

where $36\hat{\mu} = 699$, $36\hat{a}_2 = \sum_E^H Q_i$, and $180\hat{b}_2 = 182$ so that $Y = 19.1 + 2.0x$. By a similar calculation the quadratic estimation function for base Detergent II is seen to be $Y = 6.7 + 7.4x - x^2$.

4. Acknowledgment

I wish to thank Dr. A. C. Olson who conducted this experimental program for bringing this problem to my attention.

References

[1] Davies, O. L., *Design and Analysis of Industrial Experiments*, Oliver and Boyd (Great Britain), Hafner (New York), 1956, Chapter 6.

[2] Scheffé, H., *The Analysis of Variance*, John Wiley and Sons, 1959, Section 5.2.

VOL. 3, NO. 1 TECHNOMETRICS FEBRUARY, 1961

Multi-Component Systems and Structures and Their Reliability

Z. W. BIRNBAUM*, J. D. ESARY AND S. C. SAUNDERS

* *University of Washington*
and
Boeing Scientific Research Laboratories

A number of recent publications have dealt with problems of analyzing the performance of multi-component systems and evaluating their reliability. For example, a comprehensive theory of two-terminal networks was presented in [1] by Moore and Shannon who, among other results, have developed methods for obtaining highly reliable systems using components of low reliability; some of their procedures are credited to earlier work by von Neumann [2]. Several of the concepts and results of the present paper are generalizations of the corresponding concepts and results of the Moore-Shannon paper. A discussion of complex systems interpreted as Boolean functions may be found in the paper [3] by Mine.

The present study deals with general classes of systems which contain two-terminal networks and most other kinds of systems considered previously as special cases, and investigates their combinatorial properties and their reliability. These classes consist, with several variants, of systems such that the more components that perform the greater the probability that the system performs. For such systems it is shown that, if each component has reliability p and the reliability of the system is denoted by $h(p)$, then under mild additional assumptions h is an S-shaped function, i.e., its graph has the shape indicated in Fig. 3.2.4.1. Some of the consequences are these: there exists a critical value of p such that above that value the reliability of the system is greater than the reliability of a single component and below that value it is smaller; for p small the system has a reliability comparable to that of a series system, and for p large to that of a parallel system; by repeatedly iterating the system, i.e., by using replicas of the system instead of single components, one obtains systems with reliability arbitrarily close to 1 if one starts with component reliability above the critical value, but with reliability arbitrarily close to 0 if one starts below that critical value.

1. INTRODUCTION

1.1 When a very complex device is constructed, consisting of a large number of components, it is often impossible to be quite sure that it will perform the task for which it was intended. Failures of components due to causes which are hard to anticipate and practically impossible to prevent may lead to failure of the entire structure. In such situations, the best one may strive for in designing the structure is to attain a high probability that it will perform its task. It has become customary to refer to the *probability that a structure will perform* the task for which it was designed as the *reliability* of that structure. A simplifying (and possibly not quite realistic) assumption is implied in this definition of reliability: it is assumed that a structure can only either perform or fail. To emphasize this assumption, we shall sometimes use the term *dichotomic reliability*

for the probability that a structure will perform its task. It is possible to introduce and study a more general concept of reliability which accounts also for the possibility of partial performance. The present study, however, will be limited to dichotomic reliability.

1.2. A similar situation arises when, instead of a complex structure, one considers single mass-produced components. There again one can hardly expect to be sure that all components will perform and can only aim at a high probability that a component, when called upon, will perform. As before, the probability that a component will perform the function for which it is intended will be called the (dichotomic) reliability of that component.

1.3. It should be pointed out that the time element does not explicitly enter into our definition of dichotomic reliability. It may enter in what is meant by "performing" or "failing", e.g., when a structure is said to perform if it functions in a certain manner for at least 400 hours, and to fail if it breaks down before that time. But there also are practical situations where a structure is considered as performing when it is in working order for instantaneous use, and not necessarily for any length of time. Our dichotomic reliability includes the time-dependent reliability as a special case.

1.4. One of the main purposes of a mathematical theory of reliability is to develop means by which one can evaluate the reliability of a structure when the reliabilities of its components are known. The present study will be concerned with this kind of mathematical development. It will be necessary for this purpose to rephrase our intuitive concepts of structure, component, reliability, etc. in more formal language, to restate carefully our assumptions, and to introduce an appropriate mathematical apparatus.

2. Formal Description of a Structure

2.1. In order to study the relationships between the reliabilities of the components of a structure and the reliability of that structure, one has to know how performance or failure of the various components affects performance or failure of the structure. For this purpose we shall describe the state of any device, single component or complex structure, by writing 1 when it performs and 0 when it fails. If a structure consists of n components it will be called a *structure of order n*. The state of all components of such a structure will be described by a vector with n coordinates

$$\mathbf{x} = (x_1 , x_2 , \cdots , x_n)$$

where $x_i = 1$ means "ith component performs" and $x_i = 0$ means "ith component fails". All possible states of the n-tuple $\mathbf{x}$ are given by the vertices of an n-dimensional cube. Some of these 2^n vectors will make our structure perform, and others will make it fail, so that the state of the structure may be written as a function of $\mathbf{x}$, call it ϕ, which assumes the value unity for those vectors which make the structure perform and zero otherwise. The function ϕ will be called the *structure function* or, in short, the *structure*.

By analogy with terms used in the theory of circuits, any vector $\mathbf{x}$ for which $\phi(\mathbf{x}) = 1$ will be called a *path* for the structure ϕ, and any vector $\mathbf{x}$ for which $\phi(\mathbf{x}) = 0$ a *cut* for that structure.

The number of components performing when the state of all components is given by a vector $\mathbf{x}$ is clearly given by the function of $\mathbf{x}$

$$s(\mathbf{x}) = \sum_{i=1}^{n} x_i$$

which will be called *the size* of $\mathbf{x}$.

The following abbreviations and definitions will be useful:

$\mathbf{x} = (x_1, \cdots, x_n)$ (state of the performance of the components)
$\mathbf{0} = (0, 0, \cdots, 0)$ (state in which all components fail)
$\mathbf{1} = (1, 1, \cdots, 1)$ (state in which all components perform)
$\mathbf{x} \geq \mathbf{y}$ when $x_i \geq y_i$ for $i = 1, 2, \cdots, n$
$\mathbf{x} > \mathbf{y}$ when $\mathbf{x} \geq \mathbf{y}$ and $x_j > y_j$ for some $j = 1, 2, \cdots, n$

$(\mathbf{x}, x_{n+1}) = (x_1, \cdots, x_n, x_{n+1})$.

2.2. Examples of Structures

2.2.1. Parallel Components

A structure is said to consist of n components in parallel if it is so designed that it fails if and only if all components fail. This structure is described by

$$\phi(\mathbf{x}) = 1 - \prod_{i=1}^{n} (1 - x_i) = \max(x_1, x_2, \cdots, x_n)$$

and its only cut is $(0, 0, \cdots, 0) = \mathbf{0}$.

2.2.2. Components in Series

A structure which performs if and only if all its components perform is said to contain its components in series. It is described by

$$\phi(\mathbf{x}) = \prod_{i=1}^{n} x_i = \min(x_1, x_2, \cdots, x_n)$$

and the only path for ϕ is $\mathbf{1}$.

2.2.3. Switching Circuits

Consider the following circuit diagrams:

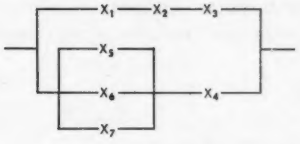


FIG. (i).

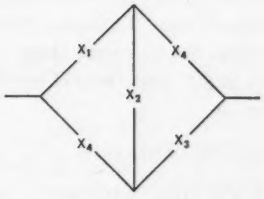


FIG. (ii).

Such diagrams as these are often used to define what we have called the structure function. By realizing that the circuit of Fig. (i) is constructed by connecting series and/or parallel subcircuits themselves in series and/or parallel we can write down the structure function using the definitions of the preceding

paragraph. Thus the structure function for the module consisting of $(x_4, \cdots, x_7)$ is

$$x_4\left[1 - \prod_{i=5}^{7}(1 - x_i)\right].$$

By continuing in this fashion we have for the entire circuit

$$\phi(x_1, \cdots, x_7) = 1 - \left(1 - \prod_{i=1}^{3} x_i\right)\left(1 - x_4\left[1 - \prod_{i=5}^{7}(1 - x_i)\right]\right). \tag{2.2.3.1}$$

We exhibit the structure function Ψ for the circuit of Fig. (ii)

$$\Psi(x_1, \cdots, x_4) = x_2[1 - (1 - x_1)(1 - x_4)][1 - (1 - x_3)(1 - x_4)] + (1 - x_2)[1 - (1 - x_3x_4)(1 - x_1x_4)]. \tag{2.2.3.2}$$

It is not our purpose to discuss in general the method of representation of switching circuits but merely to present some examples of these circuits which are structures of the type we are considering.

We call the reader's attention to the fact that a circuit of the type in Fig. (i) is sometimes called a two-terminal network. We also point out that the circuit of Fig. (ii) contains the fourth component in two different locations. This can happen for example with contacts identified in wiring or with components which receive power from the same source or again with circuit breakers connected by a rod.

2.2.4. "k out of n" structures

Let

$$\phi(\mathbf{x}) = \begin{cases} 1 & \text{if } s(\mathbf{x}) \geq k \\ 0 & \text{if } s(\mathbf{x}) < k \end{cases}$$

where k is an integer, $0 \leq k \leq n$. This structure performs when at least k of its n components perform, and fails otherwise. We shall call it a "k out of n" structure. Clearly all $\mathbf{x}$ such that $s(\mathbf{x}) \geq k$ are paths, all others cuts. For $k = 1$ the k out of n structure reduces to the structure with parallel components, and for $k = n$ to the structure with components in series.

We remark that "k out of n" structures are frequently encountered in practice, viz., it could happen that three out of four airplane engines operational during a certain flight stage would not be disastrous but the loss of one more engine would be.

2.3. The Dual Structure

2.3.1. For a given structure ϕ, we define the structure

$$\phi^D(\mathbf{x}) = 1 - \phi(1 - \mathbf{x})$$

and call it the structure *dual* to ϕ. One verifies immediately that if a vector $\mathbf{x}$ is a path for ϕ then $1 - \mathbf{x}$ is a cut for ϕ^D and vice versa, and that $(\phi^D)^D = \phi$.

2.3.2. As an example consider the k out of n structure of 2.2.3. The dual structure is

$$\phi^D(\mathbf{x}) = \begin{cases} 1 & \text{if } s(\mathbf{x}) \geq n - k + 1 \\ 0 & \text{if } s(\mathbf{x}) < n - k + 1, \end{cases}$$

so that for the k out of n structure the dual is the $n - k + 1$ out of n structure. In particular for $k = 1$ we find that the structure dual to that with parallel components is the series structure, and for $k = n$ we obtain the converse statement.

It may happen that certain structures are self dual, i.e., $\phi^D = \phi$. For example in Fig. (ii) of 2.2.3 the structure is self dual. This can be checked easily by making the appropriate changes in (2.2.3.2).

2.4. Path-numbers and Cut-numbers

A given structure ϕ may have paths and cuts of sizes ranging from 0 to n. We define the *path-numbers* for a given structure for $j = 0, 1, \cdots, n$ as

(2.4.1) $\qquad A_j =$ number of paths for ϕ of size j,

and the *cut-numbers* for ϕ for $j = 0, 1, \cdots, n$ as

(2.4.2) $\qquad A_j^* =$ number of cuts for ϕ of size j.

Let S_j be the set of all vectors of size j

(2.4.3) $$S_j = \{\mathbf{x}: s(\mathbf{x}) = j\}.$$

Then clearly

(2.4.4) $$A_j = \sum_{\mathbf{x} \in S_j} \phi(\mathbf{x}), \qquad A_j^* = \sum_{\mathbf{x} \in S_j} [1 - \phi(\mathbf{x})],$$

and since S_j contains exactly $\binom{n}{j}$ vectors and each of them is either a path or a cut we have

(2.4.5) $$A_j + A_j^* = \binom{n}{j}.$$

Setting

$$A_j^D = \text{number of paths for } \phi^D \text{ of size } j,$$

$$A_j^{D*} = \text{number of cuts for } \phi^D \text{ of size } j,$$

we obtain

(2.4.6) $$A_j^D = A_{n-j}^*, \qquad A_j^{D*} = A_{n-j}.$$

2.5. Length and Width of a Structure

By analogy with terms used in the theory of circuits we define for a structure ϕ

(2.5.1) $\qquad l =$ length of $\phi =$ smallest j such that $A_j > 0$,

(2.5.2) $\qquad w =$ width of $\phi =$ smallest k such that $A_{n-k}^* > 0$.

Intuitively speaking, l is the smallest number of components such that, if these components perform, the structure performs even if all other $n - l$ components fail. Similarly w is the smallest number of components such that if they fail the structure fails, even if all remaining $n - w$ components perform. For example, in a k out of n structure $l = k$, $w = n - k + 1$.

According to (2.5.1) it follows that

$$A_j = 0 \quad \text{for} \quad j \leq l - 1$$

hence by (2.4.5),

$$A_j^* = \binom{n}{j} > 0 \quad \text{for} \quad j \leq l - 1.$$

But also according to (2.5.1)

$$A_{n-(w-1)}^* = 0$$

hence $n - (w - 1) \geq l$ and we obtain the inequality

$$l + w \leq n + 1 \tag{2.5.3}$$

2.6. Combination of Structures

Up to now we have always considered one structure of given order n. We shall now consider structures of different orders and in particular define an operation by which structures of a higher order can be obtained from structures of a lower order.

We shall say that the structure ϕ of order $n + 1$ is a linear combination or, in short, a *combination* of the structures λ and μ of order n, when the identity holds

$$\phi(\mathbf{x}, x_{n+1}) = x_{n+1}\lambda(\mathbf{x}) + (1 - x_{n+1})\mu(\mathbf{x}). \tag{2.6.1}$$

Clearly if λ and μ are structures of order n then the right side of (2.6.1) is a structure of order $n + 1$. Conversely, every structure ϕ of order $n + 1$ can be represented as a combination of two structures of order n, i.e., can be written in the form (2.6.1), since one always has the well-known identity

$$\phi(\mathbf{x}, x_{n+1}) = x_{n+1}\phi(\mathbf{x}, 1) + (1 - x_{n+1})\phi(\mathbf{x}, 0). \tag{2.6.2}$$

This representation of a structure as a combination of structures of lower order can be carried on until one reaches structures of order one. We then have the well-known representation

$$\phi(\mathbf{x}) = \sum_{\mathbf{y}} \prod_{i=1}^{n} x_i^{y_i} (1 - x_i)^{1-y_i} \phi(\mathbf{y}) \tag{2.6.3}$$

where the sum is extended over all vectors $\mathbf{y}$ of order n.

2.7. Semi-coherent and Coherent Structures

2.7.1. Definitions

Most structures occuring in practice are so designed that if a structure performs for a state $\mathbf{x}$ of its components then it performs for every state $\mathbf{y} > \mathbf{x}$,

i.e., whenever some components which have value 0 (do not perform) in $\mathbf{x}$ are given value 1 (are made to perform). This leads us to the following definitions.

A structure ϕ is *semi-coherent* when

$$\phi(\mathbf{y}) \geq \phi(\mathbf{x}) \quad \text{for all} \quad \mathbf{y} > \mathbf{x}. \tag{2.7.1.1}$$

A structure ϕ is *coherent* if it satisfies (2.7.1.1) and

$$\phi(0) = 0, \qquad \phi(1) = 1. \tag{2.7.1.2}$$

Thus a coherent structure is semi-coherent, with the additional property that it fails when all its components fail and performs when all components perform.

For given order n, there are only two structures which are semi-coherent but not coherent: the structure $\phi(\mathbf{x}) \equiv 0$ which fails for every state of its components, and the structure $\phi(\mathbf{x}) \equiv 1$ which performs for every state of its components. To see this one only has to note that $\mathbf{0} \leq \mathbf{x} \leq \mathbf{1}$ for every $\mathbf{x}$, hence by (2.7.1.1)

$$\phi(0) \leq \phi(\mathbf{x}) \leq \phi(1),$$

and if ϕ is not coherent then either $\phi(\mathbf{0}) = 1$ hence $\phi(\mathbf{x}) \equiv 1$ or $\phi(1) = 0$ and $\phi(\mathbf{x}) \equiv 0$.

If ϕ is semi-coherent (coherent) then ϕ^D is semi-coherent (coherent).

2.7.2. Theorem

A structure ϕ of order $n + 1$ is semi-coherent if and only if it can be represented as a linear combination of functions of order n in the form (2.6.1) with λ and μ *a*) semi-coherent and b) such that $\lambda(\mathbf{x}) \geq \mu(\mathbf{x})$ for all $\mathbf{x}$. It is coherent if and only if in addition to a) and b) either c_1) $\lambda(\mathbf{x}) = \mu(\mathbf{x})$ for all $\mathbf{x}$ and λ is a coherent structure, or c_2) $\lambda(\mathbf{x}) > \mu(\mathbf{x})$ for some $\mathbf{x}$.

Proof: If ϕ is a semi-coherent structure then (2.6.2) is a representation of ϕ as a linear combination of $\lambda(\mathbf{x}) = \phi(\mathbf{x}, 1) \geq \phi(\mathbf{x}, 0) = \mu(\mathbf{x})$, hence λ and μ are semi-coherent and $\lambda(\mathbf{x}) \geq \mu(\mathbf{x})$ for every $\mathbf{x}$. If ϕ is coherent then either $\lambda(\mathbf{x}) = \phi(\mathbf{x}, 1) = \phi(\mathbf{x}, 0) = \mu(\mathbf{x})$ for every $\mathbf{x}$, and $\lambda(\mathbf{1}) = \mu(\mathbf{1}) = \phi(\mathbf{1}) = 1$, $(\lambda\mathbf{0}) = \mu(\mathbf{0}) = \phi(\mathbf{0}) = 0$ and λ and μ are coherent; or there is an $\mathbf{x}^*$ such that $\lambda(\mathbf{x}^*) > \mu(\mathbf{x}^*)$.

If ϕ can be written

$$\begin{aligned}\phi(\mathbf{x}, x_{n+1}) &= x_{n+1}\lambda(\mathbf{x}) + (1 - x_{n+1})\mu(\mathbf{x}) \\ &= x_{n+1}[\lambda(\mathbf{x}) - \mu(\mathbf{x})] + \mu(\mathbf{x})\end{aligned}$$

with λ and μ semi-coherent, then

$$\begin{aligned}\Delta &= \phi(\mathbf{y}, y_{n+1}) - \phi(\mathbf{x}, x_{n+1}) \\ &= (y_{n+1} - x_{n+1})[\lambda(\mathbf{y}) - \mu(\mathbf{y})] + x_{n+1}[\lambda(\mathbf{y}) - \lambda(\mathbf{x})] + (1 - x_{n+1})[\mu(\mathbf{y}) - \mu(\mathbf{x})].\end{aligned}$$

Assume $y_i \geq x_i$ for $i = 1, \cdots, n + 1$. If a) and b) are satisfied then either $x_{n+1} = 0$ and

$$\Delta = y_{n+1}[\lambda(\mathbf{y}) - \mu(\mathbf{y})] + [\mu(\mathbf{y}) - \mu(\mathbf{x})] = 0$$

or $1 = x_{n+1} = y_{n+1} = 1$ and $\Delta = \lambda(\mathbf{y}) - \lambda(\mathbf{x}) \geq 0$, hence ϕ is semi-coherent. If in addition c_1) holds then $\phi(\mathbf{x}, x_{n+1}) = \lambda(\mathbf{x})$ for all $(\mathbf{x}, x_{n+1})$ and ϕ is coherent. If a), b) and c_2) hold, then there is a vector $\mathbf{x}^*$ such that $1 = \lambda(\mathbf{x}^*) > \mu(\mathbf{x}^*) = 0$, hence $\phi(1) = 1, \phi(0) = 0$ and ϕ is coherent.

2.7.3. The preceding theorem suggests the following constructive procedure for obtaining all semi-coherent structures.

Let $\mathcal{C}_n$, for $n = 1, 2, \cdots$, denote the class of structures of order n defined as follows.

$\mathcal{C}_1$ consists of the three structures of order one:

$$\beta_0(x_1) = 0 \quad \text{for} \quad x_1 = 0, x_1 = 1,$$

$$\beta_1(x_1) = x_1 \quad \text{for} \quad x_1 = 0, x_1 = 1,$$

$$\beta_2(x_1) = 1 \quad \text{for} \quad x_1 = 0, x_1 = 1.$$

In concrete terms, these are one-component structures such that β_0 never performs (circuit with one contact, grounded), β_1 performs if and only if the component performs (circuit with one operative contact), β_3 always performs (circuit with one contact, shorted).

We define $\mathcal{C}_2$ as the class of all structures of order 2 of the form

$$\phi(x_1, x_2) = x_2\lambda(x_1) + (1 - x_2)\mu(x_1)$$

where $\lambda, \mu \in \mathcal{C}_1$ and $\lambda(x_1) \geq \mu(x_1)$ for $x_1 = 0, 1$.

Having defined $\mathcal{C}_n$ we define recursively $\mathcal{C}_{n+1}$ as the class of all structures of order $n + 1$ of the form

$$\phi(\mathbf{x}, x_{n+1}) = x_{n+1}\lambda(\mathbf{x}) + (1 - x_{n+1})\mu(\mathbf{x})$$

where $\lambda, \mu \in \mathcal{C}_n$ and $\lambda(\mathbf{x}) \geq \mu(\mathbf{x})$ for all $\mathbf{x}$.

Clearly $\mathcal{C}_1 \subset \mathcal{C}_2 \subset \cdots \subset \mathcal{C}_n \subset \mathcal{C}_{n+1} \subset \cdots$, and from 2.7.2 it follows that $\mathcal{C}_n$ is exactly the class of all semi-coherent structures of order n.

2.7.4. An Inequality for Path-numbers

Let ϕ be a semi-coherent structure of order n and $A_0, A_1, \cdots, A_n$ its path-numbers. Then

$$(2.7.4.1) \qquad (j + 1)A_{j+1} \geq (n - j)A_j \quad \text{for} \quad j = 0, 1, \cdots, n,$$

or equivalently,

$$(2.7.4.2) \qquad \frac{A_j}{\binom{n}{j}} \leq \frac{A_{j+1}}{\binom{n}{j+1}} \leq 1.$$

Proof: For two consecutive integers $0 \leq j, j + 1 \leq n$ we consider the sets S_j, S_{j+1} of vectors of size j and $j + 1$ and, for the purpose of this proof only,

denote the elements of $\mathcal{S}_i$ by $\mathbf{x}$ and of $\mathcal{S}_{i+1}$ by $\mathbf{y}$. For each specific $\mathbf{x}$, we denote by $\mathcal{Y}(\mathbf{x})$ the set of all $\mathbf{y}$ such that $\mathbf{y} \geq \mathbf{x}$:

$$\mathcal{Y}(\mathbf{x}) = \{\mathbf{y}: \mathbf{y} \geq \mathbf{x}\}.$$

Each set $\mathcal{Y}(\mathbf{x})$ contains $(n - j)$ different $\mathbf{y}$'s. Since for each y there are $(j + 1)$ different $\mathbf{x}$'s such that $\mathbf{y} \geq \mathbf{x}$, the collection of all elements of all sets $\mathcal{Y}(\mathbf{x})$ as $\mathbf{x}$ ranges through $\mathcal{S}_i$ forms $(j + 1)$ replicas of $\mathcal{S}_{i+1}$. Therefore

$$\sum_{\mathbf{x} \in \mathcal{S}_j} \sum_{\mathbf{y} \in \mathcal{Y}(\mathbf{x})} \phi(\mathbf{y}) = (j + 1) \sum_{\mathbf{y} \in \mathcal{S}_{j+1}} \phi(\mathbf{y}) = (j + 1) A_{j+1}$$

and since

$$\sum_{\mathbf{x} \in \mathcal{S}_j} \sum_{\mathbf{y} \in \mathcal{Y}(\mathbf{x})} \phi(\mathbf{y}) \geq \sum_{\mathbf{x} \in \mathcal{S}_j} \sum_{\mathbf{y} \in \mathcal{Y}(\mathbf{x})} \phi(\mathbf{x}) = \sum_{\mathbf{x} \in \mathcal{S}_j} \phi(\mathbf{x})(n - j) = (n - j) A_j$$

we obtain (2.7.4.1).

2.7.5. Inequality (2.7.4.1) points to the fact that the requirement of a structure being semi-coherent imposes considerable restrictions on the path-numbers A_j. Several observations are worth making in this connection.

2.7.5.1. Inequality (2.7.4.1) is necessary but not sufficient for a structure being semi-coherent, as is shown by the following example. The structure of order 3

$$\phi(0, 0, 0) = 0, \quad \phi(0, 0, 1) = 1, \quad \phi(0, 1, 0) = 1$$
$$\phi(0, 1, 1) = 1, \quad \phi(1, 0, 0) = 0, \quad \phi(1, 0, 1) = 1$$
$$\phi(1, 1, 0) = 0, \quad \phi(1, 1, 1) = 1$$

has the paths

of size 0: none
of size 1: (0, 0, 1), (0, 1, 0)
of size 2: (0, 1, 1), (1, 0, 1)
of size 3: (1, 1, 1),

hence its path-numbers are

$$A_0 = 0, \quad A_1 = 2, \quad A_2 = 2, \quad A_3 = 1$$

and (2.7.4.1) is fulfilled. But ϕ is not semi-coherent since e.g. $(1, 1, 0) > (0, 1, 0)$ and $\phi(1, 1, 0) = 0 < 1 = \phi(0, 1, 0)$.

2.7.5.2. If in the sequence of path-numbers $A_0, A_1, \cdots, A_n$ of a semi-coherent structure one has $A_j = 0$, then $A_i = 0$ for all $i \leq j$, and if one has $A_k = \binom{n}{k}$ then $A_i = \binom{n}{i}$ for all $i \geq k$. Both statements follow from (2.7.4.2).

2.7.5.3. For ϕ coherent, obviously $A_0 = 0$, $A_n = 1$. Furthermore, for the length and width of ϕ one has $l \geq 1$, $w \geq 1$, and inequality (2.5.3) becomes

(2.7.5.3) $$2 \leq l + w \leq n + 1.$$

2.7.6. Reduction of Coherent Structures

2.7.6.1. After a structure is designed, it may be possible to simplify it by

omitting components which turn out to be inessential for its performance. To do this it would be helpful to have criteria by which one can tell whether a given structure has such inessential components. A criterion of this kind will be given in 2.7.6.4, but before it can be stated we need the following definitions.

2.7.6.2. Definitions

A vector $\mathbf{z}$ is a *minimal path* for the coherent structure ϕ when $\phi(\mathbf{z}) = 1$, but for every $\mathbf{x} < \mathbf{z}$ we have $\phi(\mathbf{x}) = 0$. The ith component is *inessential* or a *dummy* component when

$$\phi(x_i , \cdots , x_{i-1} , 0, x_{i+1} , \cdots , x_n) = \phi(x_i , \cdots , x_{i-1} , 1, x_{i+1} , \cdots , x_n)$$

for all values of $x_1 , \cdots , x_{i-1} , x_{i+1} , \cdots , x_n$. Otherwise the ith component is an *essential* component.

A structure ϕ is called *irreducible* if all its components are essential.

2.7.6.3. Remarks

One verifies easily that if $\mathbf{x}$ is a path for a coherent structure then there is at least one minimal path $\mathbf{z}$ such that $\mathbf{z} \leq \mathbf{x}$.

If a structure ϕ of order n has a dummy ith component, then we have identically

$$\phi(x_1 , \cdots , x_{i-1} , x_i , x_{i+1} , \cdots , x_n) = \Psi(x_1 , \cdots , x_{i-1} , x_{i+1} , \cdots , x_n)$$

so that the ith component can be omitted and the resulting structure Ψ with only $n - 1$ components is equivalent with the original structure ϕ.

Let us now consider a structure ϕ of order $n + 1$ and assume that all its minimal paths are $\mathbf{z}^{(1)}, \mathbf{z}^{(2)}, \cdots , \mathbf{z}^{(k)}$. Each of these minimal paths is a vector

$$\mathbf{z}^{(j)} = (z_1^{(j)}, z_2^{(j)}, \cdots , z_{n+1}^{(j)}), \qquad j = 1, 2, \cdots , k,$$

with coordinates 0 or 1. We consider the vector defined by

$$\mathbf{z}^* = (\max_j z_1^{(j)}, \max_j z_2^{(j)}, \cdots , \max_j z_{n+1}^{(j)}) \tag{2.7.6.3}$$

$$= (z_1^* , z_2^* , \cdots , z_{n+1}^*).$$

The meaning of this vector is this: if $z_i^* = 1$, then the coordinate x_i assumes the value 1 for at least one of the minimal paths, that is, the ith component is required to perform for at least one minimal path; if $z_i^* = 0$, then the ith coordinate does not perform for any of the k minimal paths. In particular, if $\mathbf{z}^* = \mathbf{1}$, then every one of the $n + 1$ components of the structure must perform for some minimal path.

2.7.6.4. Theorem

Those coordinates of the vector $\mathbf{z}^*$ in (2.7.6.3) which have the value 1 correspond to essential components, and those which have the value 0 to dummy components. In particular, a necessary and sufficient condition for a coherent structure being irreducible is $\mathbf{z}^* = \mathbf{1}$.

Proof: Without loss of generality we may consider the coordinate z_{n+1}^*. If $z_{n+1}^* = 1$ then there is a minimal path, say $(\mathbf{z}, z_{n+1})$ such that $z_{n+1} = 1$. Since it is minimal and $(\mathbf{z}, 0) < (\mathbf{z}, 1)$ we have $\phi(\mathbf{z}, 1) = 1$ and $\phi(\mathbf{z}, 0) = 0$ and the $(n + 1)st$

component is essential. If $z^*_{n+1} = 0$ then the last component is zero for all minimal paths. Consider any vector $(\mathbf{x}, x_{n+1})$: it is a path or a cut for ϕ. We shall show that the $(n+1)st$ component is dummy.

If $(\mathbf{x}, x_{n+1})$ is a path then there exists a minimal path, say again $(\mathbf{z}, 0)$, and $(z, 0) \leq (\mathbf{x}, 0) \leq (\mathbf{x}, x_{n+1}) \leq (\mathbf{x}, 1)$. But $\phi(\mathbf{z}, 0) = 1$ and by monotoniety of ϕ we have $\phi(\mathbf{x}, 0) = \phi(\mathbf{x}, 1) = 1$ and the $(n+1)st$ component is dummy.

If $(\mathbf{x}, x_{n+1})$ is a cut then either $x_{n+1} = 0$ or $x_{n+1} = 1$. Take the first case: Now $(\mathbf{x}, 0)$ a cut implies $(\mathbf{x}, 1)$ is a cut also. For if it were a path the preceding argument applied would yield $1 = \phi(\mathbf{x}, 0)$ which is a contradiction to $(\mathbf{x}, 0)$ being a cut. Thus if $(\mathbf{x}, 0)$ is a cut we have $\phi(\mathbf{x}, 0) = \phi(\mathbf{x}, 1) = 0$ and the conclusion is true. Now take case two. Then $(\mathbf{x}, 1)$ is a cut but $\phi(\mathbf{x}, 0) \leq \phi(\mathbf{x}, 1) = 0$ and the proof is complete.

2.7.6.5. Remark

Theorem 2.7.6.4 suggests the following procedure for simplifying a coherent structure ϕ: first one lists all minimal paths for ϕ; then one prepares a list of all those components which must perform (have value 1) in at least one minimal path. If this list contains all n components of ϕ then this structure is irreducible. If a component does not occur in this list then that component is inessential and can be omitted. Of course the criterion of Theorem 2.7.6.4 may not be the most practical or direct way of finding inessential components. The designer should usually be able to find inessential components by direct application of engineering principles.

2.7.7. Representation of Coherent Structures

To state a simple and rather useful representation theorem, we define for any pair of vectors $\mathbf{x}$ and $\mathbf{y}$ of the same order the structure function

$$\delta(\mathbf{x}, \mathbf{y}) = \begin{cases} 1 & \text{if } \mathbf{x} \geq \mathbf{y} \\ 0 & \text{otherwise.} \end{cases}$$

Clearly, for given $\mathbf{z}$, $\delta(\mathbf{x}, \mathbf{z})$ is the coherent structure which has the minimal path $\mathbf{z}$ and no other minimal path. Similarly, for given $\mathbf{y}$, $1 - \delta(\mathbf{y}, \mathbf{x})$ is the coherent structure with the unique minimal cut $\mathbf{y}$. (A vector $\mathbf{y}$ is a minimal cut for the coherent structure ϕ when $\phi(\mathbf{y}) = 0$, but for every $\mathbf{x} > \mathbf{y}$ we have $\phi(\mathbf{x}) = 1$.)

2.7.7.1. Theorem

If ϕ is a coherent structure of order n with the minimal paths $\mathbf{z}^{(1)}, \cdots, \mathbf{z}^{(k)}$ and the minimal cuts $\mathbf{y}^{(1)}, \cdots, \mathbf{y}^{(m)}$, then it is represented by either of the expressions

$$\phi(\mathbf{x}) = 1 - \prod_{j=1}^{k} [1 - \delta(\mathbf{x}, \mathbf{z}^{(j)})], \tag{2.7.7.1.1}$$

$$\phi(\mathbf{x}) = \prod_{j=1}^{m} [1 - \delta(\mathbf{y}^{(j)}, \mathbf{x})]. \tag{2.7.7.1.2}$$

Proof: We have $\phi(\mathbf{x}) = 1$ if and only if $\mathbf{x} \geq \mathbf{z}^{(j)}$ for some j, hence if and only

if $\delta(\mathbf{x}, \mathbf{z}^{(j)}) = 1$ for some j, and this is also necessary and sufficient for the right side of (2.7.7.1.1) having the value 1. The proof of (2.7.7.1.2) is similar.

2.7.7.2. Remarks

The identities of 2.7.7.1, together with the obvious identity

$$\delta(\mathbf{x}, \mathbf{z}) = \prod_{(i\,:\,z_i=1)} x_i \quad \text{for} \quad \mathbf{z} \neq 0, \tag{2.7.7.2.1}$$

make it possible to write down immediately the structure function of a structure for which either the minimal paths or the minimal cuts are given. It may also be worthwhile to note a similarity between the identities of 2.7.7.1 and those in 2.2.1 and 2.2.2, and to observe that the structure ϕ behaves as if it consisted of the structures $\delta(\mathbf{x}, \mathbf{z}^{(j)})$ arranged in parallel, or of the structures $1 - \delta(\mathbf{y}^{(j)}, \mathbf{x})$ arranged in series.

2.7.7.3. Example

Let us take the structure defined by Fig. (ii) of (2.2.3) for which a representation is given in (2.2.3.2). It has minimal paths $\mathbf{z}^{(1)} = (1, 0, 0, 1)$, $\mathbf{z}^{(2)} = (0, 1, 0, 1)$, $\mathbf{z}^{(3)} = (0, 0, 1, 1)$, $\mathbf{z}^{(4)} = (1, 1, 1, 0)$. Thus we see

$$\Psi(x_1, \cdots, x_4) = 1 - (1 - x_1x_4)(1 - x_2x_4)(1 - x_3x_4)(1 - x_1x_2x_3)$$

which may be compared with that given earlier. However from this representation it is not easy to see that the structure is self dual.

2.8. Composition of Structures

In practical designing of structures it often happens that the design is prepared in stages, so that first a structure of order m is conceived, and then for some or all of its m components one substitutes other multi-component structures. This is for example the case when in planning electric circuits self-contained packages of components called "modules" are used as "components" of a circuit. The formal description of this process of superimposing structures is introduced by the following definition.

Let γ be a structure of order n, and Ψ_i be a structure of order k_i for $i = 1, \cdots, n$. Then the structure χ of order not exceeding $\sum_1^n k_i$ defined by

$$\chi(\mathbf{x}, \mathbf{y}, \cdots, \mathbf{z}) = \gamma(\Psi_1(\mathbf{x}), \Psi_2(\mathbf{y}), \cdots, \Psi_n(\mathbf{z}))$$

is called the *composition of* $\Psi_1, \cdots, \Psi_n$ *into* γ, where $\mathbf{x}, \mathbf{y}, \cdots, \mathbf{z}$ are vectors of order $k_1, \cdots, k_n$ respectively.

The blocks of components $\mathbf{x}, \mathbf{y}, \cdots, \mathbf{z}$ may overlap. In particular, all of them may consist of the same m components, say $\mathbf{x} = (x_1, \cdots, x_m)$, so that the resulting structure

$$\chi(x_1, \cdots, x_m) = \gamma(\Psi_1(\mathbf{x}), \cdots, \Psi_m(\mathbf{x}))$$

is of order m.

Both of the switching circuits defined in (2.2.3) may be obtained by the

composition of simpler structures to make the more complex ones indicated. It is in fact by this method that we exhibited the structure function of the circuit diagram given.

One verifies that if γ and $\Psi_1 , \cdots , \Psi_n$ are all semi-coherent then χ is semi-coherent, and if $\gamma, \Psi_1 , \cdots , \Psi_n$ are coherent then χ is coherent.

2.8.2. Combination of structures, as defined in 2.6, is a special case of composition, since (2.6.1) can be obtained by the composition of β_1 , λ, μ into the structure γ defined by

$$\gamma(u, v, w) = uw + (1 - u)w \tag{2.8.2.1}$$

where λ, μ have the same block of components $\mathbf{x}$, and β_1 the component x_{n+1}. One can therefore, beginning with the structure γ of order 3 given in (2.8.2.1) which is coherent, and using semi-coherent λ and μ of increasing orders, obtain exactly the same family of structures which can be obtained beginning with $\beta_0 , \beta_1 , \beta_2$ of 2.7.3 and proceeding by successive combinations.

3. The Reliability Function

3.1. Definition of Reliability Function

3.1.1. As indicated in section 1, one of the main aims of the mathematical theory of reliability is to evaluate the probability that a given structure will perform. Having previously discussed formal properties of structures, we shall now assume that components of a given structure may perform or fail in a random manner and derive various statements about the reliability of the structure.

3.1.2. Let ϕ be a structure of order n. We assume that its n components are independent random variables

$$\mathbf{X} = (X_1 , X_2 , \cdots , X_n),$$

each with the same probability distribution

$$\Pr \{X_i = 1\} = p, \qquad \Pr \{X_i = 0\} = 1 - p, \tag{3.1.2.1}$$

so that

$$\phi(\mathbf{X}) = \phi(X_1 , \cdots , X_n)$$

is also a random variable capable of the values 1 and 0. We now define the *reliability function* of the structure ϕ as

$$h_\phi(p) = \Pr \{\phi(\mathbf{X}) = 1\} = E[\phi(\mathbf{X})], \tag{3.1.2.2}$$

and devote the remainder of this paper to the study of reliability functions.

3.2. General Properties of Reliability Functions

3.2.1. From (3.1.2.1), (3.1.2.2) and (2.4.1) follows immediately

$$h_\phi(p) = \sum_{i=0}^{n} A_i p^i (1 - p)^{n-i}. \tag{3.2.1}$$

From this it is clear that $h_\phi(0) = 0$ if and only if $A_0 = 0$ and $h_\phi(1) = 1$ if and only if $A_n = 1$.

3.2.2. If the structure ϕ of order $n + 1$ is a combination of structures λ and μ of order n, then by taking mathematical expectations on both sides of (2.6.1) and making use of the independence of $X_1, X_2, \cdots, X_{n+1}$ one obtains

$$h_\phi(p) = p h_\lambda(p) + (1 - p) h_\mu(p). \tag{3.2.2}$$

3.2.3. We now consider structures obtained by composition in the following manner. Let ϕ be a structure of order n, Ψ a structure of order m, and let $(Y_1, \cdots, Y_{mn})$ be independent random variables, all with the same probability distribution

$$\Pr\{Y_i = 1\} = p, \qquad \Pr\{Y_i = 0\} = 1 - p.$$

Then the structure χ of order nm defined by $\chi(Y_1, Y_2, \cdots, Y_{nm})$ equal to

$$\phi[\Psi(Y_1, \cdots, Y_m), \quad \Psi(Y_{m+1}, \cdots, Y_{2m}), \cdots, \Psi(Y_{(n-1)m+1}, \cdots, Y_{nm})]$$

has the reliability function

$$h_\chi(p) = h_\phi(h_\Psi(p)). \tag{3.2.3}$$

3.2.4. The reliability functions of most practical structures have the following plausible qualitative properties:

a) $h(0) = 0$ and $h(1) = 1$,
b) $h'(p) > 0$ for $0 < p < 1$,
c) $h(p) < p$ in some neighborhood of 0
$h(p) > p$ in some neighborhood of 1 and there is exactly one root of the equation $h(p) = p$ in the open interval $(0, 1)$.

Necessary and sufficient conditions for ϕ to have h_ϕ with property a) were stated in 3.2.1. We shall also obtain conditions on a structure ϕ which are necessary and sufficient in order that h_ϕ have property b), or a more precisely defined property c). Before doing this, however, some intuitive comments on these properties may be in order.

Property b) tells that the reliability of the structure increases as the component reliability p increases, and is encountered for practically all structures.

When h has property c) we will say that h is *S-shaped*, since its graph has then the general form indicated in Fig. 3.2.4.1. A formal theory of S-shaped functions will be presented later, but already in this qualitative discussion we may indicate some consequences of h being S-shaped.

If h is S-shaped and equation $h(p) = p$ has the root $0 < \hat{p} < 1$, then we have $h(p) < p$ and the structure is less reliable than any single component as long as $p < \hat{p}$, and similarly $h(p) > p$ and the structure is more reliable than any single component when $p > \hat{p}$.

Furthermore, let ϕ be a structure with an S-shaped reliability function h, and let the sequence of *iterated compositions* of ϕ be defined as the sequence of structures

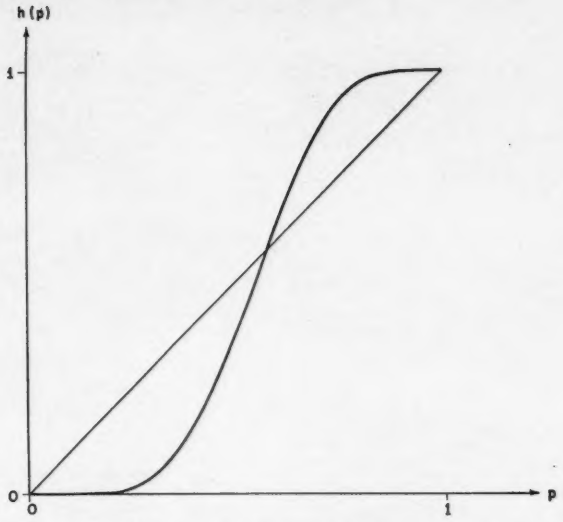


FIGURE 1—S-shaped function.
(reliability function for a 6 out of 10 structure)

$$\phi_1(x_1, \cdots, x_n) = \phi(x_1, \cdots, x_n)$$

$$\phi_2(x_1, \cdots, x_{n^2}) = \phi(\phi_1(x_1, \cdots, x_n), \cdots, \phi_1(x_{n^2-n+1}, \cdots, x_{n^2}))$$

$$\cdots$$

$$\phi_k(x_1, \cdots, x_{n^k}) = \phi(\phi_{k-1}(x_1, \cdots, x_{n^{k-1}}), \cdots, \phi_{k-1}(x_{n^k-n^{k-1}+1}, \cdots, x_{n^k}))$$

$$\cdots .$$

According to 3.2.3 the corresponding reliability functions are

$$h_1(p) = h(p)$$

$$h_2(p) = h[h_1(p)]$$

$$\cdots \qquad \cdots$$

$$h_k(p) = h[h_{k-1}(p)]$$

$$\cdots \qquad \cdots .$$

Let again $h(\hat{p}) = \hat{p}$, $0 < \hat{p} < 1$. Then it can be seen by an argument indicated by Fig. 3.2.4.2 that

$$h_1(p) > h_2(p) > \cdots > h_k(p) \to 0 \quad \text{if} \quad p < \hat{p},$$

$$h_1(p) < h_2(p) < \cdots < h_k(p) \to 1 \quad \text{if} \quad p > \hat{p},$$

so that by iterated compositions of ϕ one can obtain structures with reliabilities as close to 1 as desired if the initial component reliability is $p > \hat{p}$, but tending to 0 if $p < \hat{p}$.

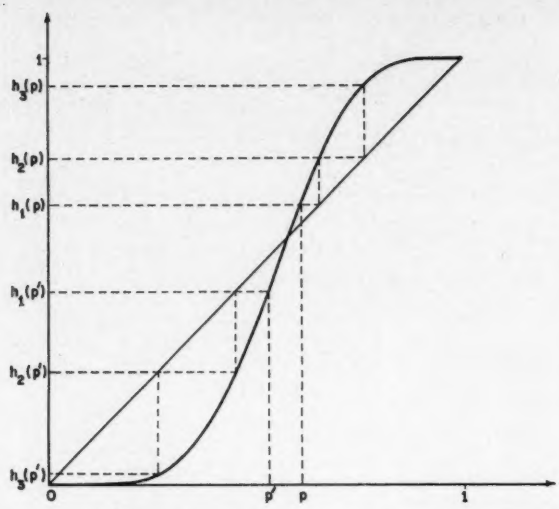


FIGURE 2—Iterated composition of reliability function for 6 out of 10 structure.

3.2.5. Example: k out of n structures

It is easily seen that the structure described in 2.2.3 has the reliability function

$$h(p; k, n) = \sum_{i=k}^{n} \binom{n}{i} p^i (1-p)^{n-i}. \tag{3.2.5.1}$$

For $n = 1, 2, \cdots, 25$, the roots $\hat{p}_{k,n}$ of the equation

$$h(p; k, n) = p, \qquad 0 < p < 1, \tag{3.2.5.2}$$

may be read off from Table 3.2.5. This table was computed from [4] by linear interpolation, except for the case $k = 2$ for which (3.2.5.2) was solved directly and the solution rounded off to three places.

For large values of n, and np, $n(1-p)$ both not small, the central limit theorem makes it possible to replace (3.2.5.1) by the approximate equality

$$h(p; k, n) \approx 1 - N(u)$$

where $N(u)$ is the area under the standard normal probability curve from $-\infty$ to u and

$$u = (k - np)/\sqrt{np(1-p)}.$$

3.3. Theory of S-shaped Functions

3.3.1. Definitions

3.3.1.1. We shall say that the function f belongs to class $\bar{\Psi}$ when it satisfies the following conditions

(i) $f(p)$ is continuous and $0 \leq f(p) \leq 1$ for $0 \leq p \leq 1$ and $f(p)$ is differentiable for $0 < p < 1$,

(ii) the function σ_f defined by

$$\sigma_f(p) = \frac{f(p)}{p} \cdot \frac{1-p}{1-f(p)}$$

is non-decreasing for $0 < p < 1$.

3.3.1.2. We say that f belongs to class Ψ when it satisfies (i) and
(ii') σ_f is strictly increasing for $0 < p < 1$,
(iii) $f(0) = 0, f(1) = 1$.

3.3.1.3. We say that f belongs to class Ψ_0 when it belongs to Ψ and satisfies the condition

(iv) $\sigma_f(p)$ assumes the value 1 for some p, $0 < p < 1$.

3.3.2. Remarks

From the definition of σ_f follows

$$f = \frac{p\sigma_f}{1-p+p\sigma_f} = \frac{1}{1+\frac{1-p}{p\sigma_f}}, \tag{3.3.2.1}$$

TABLE 3.2.5

Solutions $\hat{p}_{k,n}$ of the equation $h(p; k, n) = p$ for $p \in (0, 1)$

k / n	2	3	4	5	6	7	8	9	10	11	12	13
3	.500											
4	.232											
5	.131	.500										
6	.083	.347										
7	.058	.256	.500									
8	.042	.197	.395									
9	.032	.158	.322	.500								
10	.025	.129	.268	.421								
11	.020	.108	.228	.361	.500							
12	.016	.092	.197	.314	.437							
13	.013	.080	.172	.276	.387	.500						
14	.011	.070	.152	.246	.345	.448						
15	.010	.062	.136	.220	.311	.405	.500					
16	.010	.055	.122	.199	.282	.368	.456					
17	.008	.050	.111	.181	.257	.337	.418	.500				
18	.007	.045	.101	.166	.236	.310	.385	.462				
19	.006	.041	.092	.153	.218	.286	.356	.428	.500			
20	.006	.037	.085	.141	.202	.266	.331	.398	.466			
21	.005	.034	.079	.131	.188	.247	.309	.372	.436	.500		
22	.005	.032	.073	.122	.175	.231	.289	.349	.409	.470		
23	.004	.030	.068	.114	.164	.217	.272	.328	.385	.442	.500	
24	.004	.027	.064	.107	.154	.204	.256	.309	.363	.418	.473	
25	.004	.025	.060	.101	.145	.193	.242	.292	.343	.395	.448	.500

For $k > n/2$:

$$\hat{p}_{k,n} = 1 - \hat{p}_{n+1-k,n}.$$

hence a function $f \in \bar{\Psi}$ is strictly increasing for every p such that $0 < \sigma_f(p) < \infty$.

Clearly $\sigma_f(p) = 1$ for exactly those p for which $f(p) = p$. Since for $f \in \Psi$ the function σ_f assumes every value at most once, it follows that Ψ_0 consists of exactly those $f \in \Psi$ for which the equation $f(p) = p$ has a unique root in $0 < p < 1$.

Let $f \in \Psi_0$ and $f(p^*) = p^*$, $0 < p^* < 1$, hence $\sigma_f(p) < 1$ for $p < p^*$, $\sigma_f(p) > 1$ for $p > p^*$. From (3.3.2.1) we see that $f(p) < p$ for $p < p^*$, $f(p) > p$ for $p > p^*$. The functions of class Ψ_0 are therefore S-shaped in the sense of 3.2.4.

Let now $f \in \bar{\Psi}$. If $f(0) > 0$ then $\sigma_f(0) = +\infty$, hence $\sigma_f(p) = +\infty$ for $0 \leq p \leq 1$ and according to (3.3.2.1) $f(p) \equiv 1$ for $0 \leq p \leq 1$. If $f(1) < 1$ then $\sigma_f(1) = 0$, hence $\sigma_f(p) = 0$ and $f(p) = 0$ for $0 \leq p \leq 1$. A function $f \in \bar{\Psi}$ therefore is either identically 0, or identically 1, or it maps the closed interval [0, 1] onto the closed interval [0, 1]. This mapping need not be one-to-one, as shown by the example

$$f(p) = \begin{cases} 0 & \text{for } 0 \leq p \leq 1/2 \\ 4(p - 1/2)^2 & \text{for } 1/2 \leq p \leq 1. \end{cases}$$

3.3.3. Theorem

Let $f, g \in \bar{\Psi}$ and at least one of f, g be neither identically 0 nor identically 1. Furthermore, let either

a) $f(p) \geq g(p)$ for $0 < p < 1$ and at least one of f, g be in Ψ or
b) $f(p) > g(p)$ for $0 < p < 1$.

Then the function h defined by

$$h(p) = pf(p) + (1 - p)g(p) \tag{3.3.3.1}$$

is in Ψ.

Proof: One verifies directly that h satisfies (i) and (iii). To prove (ii'), that is $\sigma_h'(p) > 0$ for $0 < p < 1$, we shall prove the equivalent statement

$$p(1 - p)h' > h(1 - h) \quad \text{for} \quad 0 < p < 1. \tag{3.3.3.2}$$

From (3.3.3.1) follows

$$p(1 - p)h' = pp(1 - p)f' + (f - g)p(1 - p) + (1 - p)p(1 - p)g',$$

$$h(1 - h) = pf(1 - f) + p(1 - p)(f - g)^2 + (1 - p)g(1 - g).$$

Since $f, g \in \bar{\Psi}$ we have $p(1 - p)f' \geq f(1 - f)$, $p(1 - p)g' \geq g(1 - g)$, and in case a) at least one of these inequalities is strict, and we have $0 \leq f - g \leq 1$ hence $f - g \geq (f - g)^2$. In case b) we have $f - g > (f - g)^2$ for $0 < p < 1$, so that (3.3.3.2) holds in either case.

3.3.4. Theorem

If $f, g \in \Psi$, then $h = f[g] \in \Psi$.

Proof: Properties (i) and (iii) for h are obvious, and (ii') follows from

$$\sigma_h(p) = \sigma_f[g(p)] \cdot \sigma_g(p).$$

3.4. Mean Path and Mean Cut

3.4.1. Let $\mathbf{X} = (X_1, X_2, \cdots, X_n)$ and $\Phi = \phi(\mathbf{X})$ have the same meaning as in 3.1.2, and let

$$S = s(\mathbf{X}) = X_1 + X_2 + \cdots + X_n$$

be the size of the random vector $\mathbf{X}$, i.e., the number of components which perform. Then S is a random variable with the binomial distribution, Φ a random variable with the probability distribution $Pr\{\Phi = 1\} = h(p)$, $Pr\{\Phi = 0\} = 1 - h(p)$, but S and Φ are dependent random variables. It is natural to consider the conditional expectations

$$E(S \mid \Phi = 1) = L(p),$$
$$E(n - S \mid \Phi = 0) = W(p)$$

and to call $L(p)$ the *mean path*, $W(p)$ the *mean cut* for the structure ϕ.

3.4.2. Theorem

In order that the reliability function h of a structure of order n is strictly increasing for $0 < p < 1$ it is necessary and sufficient that

(3.4.2.1) $$L(p) + W(p) > n.$$

This condition is equivalent with

(3.4.2.2) $$\operatorname{cov}(S, \Phi) > 0.$$

Proof: We shall prove somewhat more specifically that for any fixed p, $0 < p < 1$, each of the conditions (3.4.2.1), (3.4.2.2) is necessary and sufficient to have $h'_\phi(p) > 0$.

Writing h in short for h_ϕ we have from (3.2.1)

$$h'(p) = \frac{1}{p(1-p)}\left[\sum_{i=0}^{n} jA_i p^i(1-p)^{n-i} - np\sum_{i=0}^{n} A_i p^i(1-p)^{n-i}\right]$$

(3.4.2.3) $$= \frac{1}{p(1-p)}\left[E(S\Phi) - E(S)E(\Phi)\right]$$

$$= \frac{1}{p(1-p)}\operatorname{cov}(S, \Phi)$$

which shows that (3.4.2.2) is equivalent with $h'(p) > 0$.

Since

$$E(S \mid \Phi = 1) = \sum_{i=0}^{n} j\frac{\Pr(S = j, \Phi = 1)}{\Pr(\Phi = 1)}$$
$$= \frac{1}{E(\Phi)}\sum_{i=0}^{n} jA_i p^i(1-p)^{n-i} = \frac{E(\Phi S)}{E(\Phi)}$$

and

$$E(S \mid \Phi = 0) = \sum_{i=0}^{n} j\frac{\Pr(S = j, \Phi = 0)}{\Pr(\Phi = 0)}$$
$$= \frac{1}{1 - E(\Phi)}\sum_{i=0}^{n} jA_i^* p^i(1-p)^{n-i}$$
$$= \frac{E(S) - E(\Phi S)}{1 - E(\Phi)},$$

we obtain

$$L(p) + W(p) = \frac{E(\Phi S)}{E(\Phi)} + n - \frac{E(S) - E(\Phi S)}{1 - E(\Phi)} \tag{3.4.2.4}$$

$$= n + \frac{\operatorname{cov}(\Phi, S)}{h(p)[1 - h(p)]}$$

so that (3.4.2.1) and (3.4.2.2) are equivalent.

3.4.3. Theorem

Let ϕ be a structure of order n and h its reliability function. In order that σ_h is strictly increasing it is necessary and sufficient that

$$L(p) + W(p) > n + 1 \tag{3.4.3.1}$$

and this condition is equivalent with

$$\operatorname{cov}(S - \Phi, \Phi) > 0. \tag{3.4.3.2}$$

Proof: We have seen in 3.3.3 that $\sigma_h'(p) > 0$ at a point p, $0 < p < 1$, is equivalent with inequality (3.3.3.2). From (3.4.2.3) follows

$$p(1 - p)h'(p) = E(S\Phi) - E(S)E(\Phi) = \operatorname{cov}(S, \Phi)$$

and since

$$h(p)[1 - h(p)] = \operatorname{var}(\Phi)$$

inequality (3.3.3.2) is equivalent with

$$\operatorname{cov}(S - \Phi, \Phi) = \operatorname{cov}(S, \Phi) - \operatorname{var}(\Phi) > 0.$$

From (3.4.2.4) follows

$$L(p) + W(p) = n + 1 + \frac{\operatorname{cov}(S - \Phi, \Phi)}{h(p)[1 - h(p)]}$$

which shows that (3.4.3.1) and (3.4.3.2) are equivalent.

3.5. Structures Coherent in Probability

3.4.1. We shall say that a structure ϕ of order n is *coherent in probability* when

(a) $\Pr\{\Phi = 1 \mid S = k\} \leq \Pr\{\Phi = 1 \mid S = k + 1\}$

for $k = 0, 1, \cdots, n - 1$,

(b) $\Pr\{\Phi = 1 \mid S = 0\} = 0, \quad \Pr\{\Phi = 1 \mid S = n\} = 1.$

3.5.2. Theorem

A structure is coherent in probability if and only if its path-numbers satisfy the inequalities (2.7.4.2) and (2.7.4.1), and $A_0 = 0$, $A_1 = 1$.

Proof: Since $Pr\{\Phi = 1, S = k\} = A_k p^k (1 - p)^{n-k}$ and

$$Pr\{S = k\} = \binom{n}{k} p^k (1 - p)^{n-k},$$

we have

$$\Pr\{\Phi = 1 \mid S = k\} = A_k \Big/ \binom{n}{k}, \qquad k = 0, 1, \cdots, n$$

for any structure, and our theorem follows immediately.

3.5.3. Corollary

If a structure is coherent, then it is coherent in probability.

This follows from 3.5.2 and 2.7.4. The converse statement is not true, as shown by the example of 2.7.5.1. The structure of order 3 described there satisfies (2.7.4.2) hence is coherent in probability, but is not even semi-coherent.

3.5.4. Theorem

If ϕ is coherent in probability then its reliability function h is non-decreasing for $0 \leq p \leq 1$.

Proof: From (3.2.1) we have

$$h'(p) = \sum_{i=0}^{n-1} [(j+1)A_{i+1} - (n-j)A_i]p^i(1-p)^{n-i-1} \tag{3.5.4.1}$$

and by 3.5.2 and (2.7.4.1) we obtain $h'(p) \geq 0$.

3.6. Reliability Functions of Semi-Coherent and Coherent Structures

3.6.1. Theorem

If ϕ is a semi-coherent structure then either $h_\phi \equiv 0$, or $h_\phi \equiv 1$, or $h_\phi \equiv p$, or $h_\phi \in \Psi$.

Proof: To use induction on the order of ϕ, we first consider $n = 1$. The three possible structures β_0, β_1, β_2 of order 1 listed in 2.7.3 have the reliability functions 0, p, 1, respectively. We now assume our statement to be true for order n, and consider a semi-coherent structure ϕ of order $n+1$. According to 2.7.2 we have

$$\phi(\mathbf{X}, X_{n+1}) = X_{n+1}\lambda(\mathbf{X}) + (1 - X_{n+1})\mu(\mathbf{X}) \tag{3.6.1.1}$$

with λ and μ semi-coherent of order n and $\lambda(\mathbf{x}) \geq \mu(\mathbf{x})$ for all $\mathbf{x}$. From (3.2.2) we have

$$h_\phi(p) = ph_\lambda(p) + (1-p)h_\mu(p).$$

If $\lambda(\mathbf{x}) = \mu(\mathbf{x})$ for all $\mathbf{x}$, then $h_\phi(p) = h_\lambda(p)$ for $0 \leq p \leq 1$ and our statement is true by the assumption of induction. If $\lambda(\mathbf{x}) > \mu(\mathbf{x})$ for some value of $\mathbf{x}$, then $h_\lambda(p) > h_\mu(p)$ for $0 < p < 1$ and $h_\phi \in \Psi$ according to 3.3.3.

3.6.2. Corollary

If ϕ is a coherent structure then either

$$h_\phi(p) \equiv p \quad \text{or} \quad h_\phi \in \Psi.$$

This follows immediately from 3.6.1 by observing that for ϕ coherent $h(0) = 0$ and $h(1) = 1$ hence h is neither identically 0 nor identically 1.

3.6.3. Theorem

If ϕ is a coherent structure, then $h_\phi \in \Psi_0$ if and only if $A_1 = 0$ and $A_{n-1} = n$.

Proof: Since for ϕ coherent $A_0 = 0$, $A_n = 1$ we have

$$\sigma_h(p) = \frac{\sum_{i=1}^{n} A_i p^i (1-p)^{n-i}}{p} \cdot \frac{1-p}{\sum_{i=0}^{n-1} \left[\binom{n}{i} - A_i\right] p^i (1-p)^{n-i}}$$

$$= \frac{\sum_{i=1}^{n} A_i p^{i-1} (1-p)^{n-i}}{\sum_{i=0}^{n-1} \left[\binom{n}{i} - A_i\right] p^i (1-p)^{n-1-i}}$$

and

$$\lim_{p \to 0} \sigma_h(p) = \frac{A_1}{1 - A_0} = A_1$$

(3.6.3.1)

$$\lim_{p \to 1} \sigma_h(p) = \frac{A_n}{\binom{n}{n-1} - A_{n-1}} = \frac{1}{n - A_{n-1}}.$$

For $h_\phi \in \Psi$ the function $\sigma_h(p)$ is strictly increasing and it assumes the value 1 in $0 < p < 1$ if and only if $\lim_{p \to 0} \sigma_h(p) < 1$ and $\lim_{p \to 1} \sigma_h(p) > 1$ which together with (3.6.3.1) shows that $A_1 = 0$, $A_{n-1} = n$.

3.6.4. Inequalities for $h(p)$ in Terms of Length and Width

Let l and w denote the length and the width of a structure ϕ, as defined in 2.5. Then $A_i = 0$ for $i = 0, 1, \cdots, l-1$, and $A_1^* = 0$ for $i = n - w + 1, \cdots, n$, so that the reliability function h of ϕ can be written

$$h(p) = \sum_{i=l}^{n} A_i p^i (1-p)^{n-i} = 1 - \sum_{i=0}^{n-w} A_i^* p^i (1-p)^{n-i}. \tag{3.6.4.1}$$

This immediately yields the inequalities

$$A_l p^l (1-p)^{n-l} \leq h(p) \leq 1 - A_{n-w}^* p^{n-w} (1-p)^w \tag{3.6.4.2}$$

which show that for $p \to 0$ the function $h(p)$ tends to 0 not faster than p^l, that is, not faster than in the case of components in series; and for $p \to 1$ it tends to 1 not faster than $1 - (1-p)^w$, that is, not faster than in the case of w parallel components.

Inequalities (3.6.4.1) hold without any assumptions on ϕ. If ϕ is coherent in probability, we have inequalities (2.7.4.2) from which one obtains

$$A_i \geq \frac{\binom{n}{i}}{\binom{n}{l}} A_l \quad \text{for} \quad i = l, l+1, \cdots, n$$

and

$$A_i^* \geq \frac{\binom{n}{i}}{\binom{n}{w}} A_{n-w}^* \quad \text{for} \quad i = 0, 1, \cdots, n - w.$$

From these inequalities and (3.6.4.1) follow the bounds for $h(p)$

$$\frac{A_l}{\binom{n}{l}} \sum_{i=l}^{n} \binom{n}{i} p^i (1-p)^{n-i} \leq h(p) \leq 1 - \frac{A_{n-w}^*}{\binom{n}{w}} \sum_{i=0}^{n-w} \binom{n}{i} p^i (1-p)^{n-i}. \tag{3.6.4.3}$$

3.6.5. Qualitative Remarks

In 3.2.4 we described some properties of reliability functions which appeared desirable for practical structures. The preceding theorems tell us what kind of structures have these properties: If ϕ is coherent then h_ϕ has properties a) and b) of 3.2.4 (according to 3.6.2). Furthermore, if ϕ is coherent then the conditions on path-numbers $A_1 = 0$, $A_{n-1}^* = n$ characterize it as having a reliability function h_ϕ of class Ψ_0, hence S-shaped and lending itself to the process of iterated compositions.

The class of coherent structures is very large. It contains among others the two-terminal networks whose reliability has been studied by Moore and Shannon in their fundamental paper [1] and all k out of n structures.

References

[1] E. F. Moore and C. E. Shannon, "Reliable circuits using less reliable relays", *Journ. of the Franklin Institute*, v. 262 (1956), pp. 191–208.

[2] J. von Neumann, "Probabilistic logics and the synthesis of reliable organisms from unreliable components," Automata Studies, pp. 43–98, *Annals of Math. Studies No. 34*, Princeton University Press 1956.

[3] Hisashi Mine, "Reliability of physical system", *IRE Transactions, PGIT*, v. IT-5, Special Supplement May 1959, pp. 138–150.

[4] *Tables of the Binomial Probability Distribution*, National Bureau of Standards, Applied Mathematics Series 6, 1950.

VOL. 3, NO. 1 TECHNOMETRICS FEBRUARY, 1961

An Asymptotic Distribution for an Occupancy Problem with Statistical Applications

M. HALPERIN AND G. L. BURROWS

Knolls Atomic Power Laboratory,

Given a population of N items D of which are defective. Let the population be divided at random into k lots of equal size s (*i.e.* $N = ks$). This paper derives the asymptotic distribution ($k \to \infty$) of the number of lots containing x or more defectives ($x = 0, 1, 2, \cdots, s$). Now suppose it is reasonable to define a lot as defective if it contains x or more defectives, and one is interested in the proportion of lots in a grand lot which are defective. Then on the basis of a random sample of items one can obtain confidence limits on various functions of the distribution of the number of defective lots. This allows high power against an excessive number of defective lots in an average sense (as in variables sampling) even for rather small samples. If, in fact, allocation of items to lots is not random but occurs with local positive contagion it appears that the random sample will tend to overstate the proportion of defective lots.

1. INTRODUCTION

Sampling is often carried out in order to estimate or put confidence limits on a probability. On a manufactured product that probability is often simply the percent of defective items produced, if defective items occur randonly. If however, the product is made, purchased or used in groups or in such a way that defectives tend to group together, the percent or number of lots or probability of a lot containing x or more defectives may be the quantities of interest. If defectives occur randomly, the probability distribution of the number of lots containing x or more defectives will be a function of the probability of a defective.

Thus, for example, a low percentage of silicon inclusions per unit volume in steel castings is sometimes tolerated in the interest of economy and for lack of 100% foolproof detection. A frequency of inclusions that is safe for small, easily replaced die castings may be quite unacceptable for huge, expensive turbine-generator rotors because of the high percentage of rotors containing at least one inclusion. Similarly, of two manufacturing processes, each producing the same proportion of defective items, one may be judged acceptable simply because it is in statistical control whereas the second may be producing an abnormally high percentage of lots each containing a substantial number of defectives. The latter process by virtue of the non-random occurrence of defectives should be correctible and may therefore be unacceptable as is.

If the main interest were the probability that m or more lots of size s, from among k such lots, contained x or more defective items, it might be quite uneconomical or difficult to select a sample of lots and ascertain for each sampled lot the number of defective items. Either of two alternatives may be available: (1) Select a subsample of individual items from each sample lot or (2) select a

sample of individual items from among all items and group the remaining items at random into lots, each of size s. We consider the second alternative.

This paper gives the exact and asymptotic distributions of $m(x)$, the number of sublots, each of size s, containing x or more defectives from among ks items containing D defectives, where the sampling of items and formation of lots are as in (2) above. Since the cumulative probability function of $m(x)$, $Pr\{m(x)\}$, is a monotone decreasing function of D, the usual upper confidence limit on D leads to a lower confidence limit on $Pr\{m(x)\}$. If the lots cannot be composed at random and if there is positive contagion, the probabilities obtained assuming random composition of lots will, under plausible assumptions as to the effect of the contagion, provide upper bounds to the true probabilities. For example, fuel plates for reactors are often made by rolling fuel ingots flat and cutting them into sections, each of which is completely covered by a protective cladding. If, in the cladding operation, the reduction of sections is slight, any fraction of a plate core volume corresponding to a particular square inch of plate surface that is known to contain high fuel content due to over-nominal thickness is apt to be contiguous to square inches also having high fuel content. The proportion of over-nominal thickness inches of rolled fuel ingots will then lead to an upper bound to the number of plates, or at least to the number of sections, with one or more square inch with over-nominal uranium loading. Thus a bound on the number of defective sections can be obtained without sampling sections; it will, of course, be a satisfactory bound only if the fuel ingots are of sufficiently uniform thickness. The distribution of $m(x)$ is also a useful guide as to whether a process is good enough to allow random assembly, as opposed to selective assembly, of its output into lots.

2. Summary

Consider a population of N like items, D of which ($0 \leq D \leq N$) are defective. Suppose we allocate the N items to $[N/s]$ groups of size s at random ($[a]$ is the usual symbol for the largest integer contained in a). Define $m(x)$, $x = 1, 2, \cdots, s$, as the observed number of groups containing x or more defectives each. It is shown that any $(s - 1)$ of the $m(x)$ have asymptotically ($N \to \infty$, $D/N \to q$, $0 < q < 1$) a multivariate normal distribution. Taking $N/s = k$ and

$$I_q(x, s - x + 1) = \frac{\int_0^q t^{x-1}(1 - t)^{s-x}\,dt}{\int_0^1 t^{x-1}(1 - t)^{s-x}\,dt},$$

then asymptotically for $0 < x \leq s$

$$E\{m(x)\} = kI_q(x, s - x + 1)$$

and

$$\begin{aligned}\text{Var}\,\{m(x)\} = {} & kI_q(x, s - x + 1)[1 - I_q(x, s - x + 1)] \\ & - ks(1 - q)I_q^2(x, s - x) - ksqI_q^2(x - 1, s - x + 1) \\ & + ksI_q^2(x, s - x + 1),\end{aligned}$$

where

$$I_q(a, b) = 0 \quad \text{if} \quad b \leq 0$$
$$= 1 \quad \text{if} \quad a \leq 0$$

$$\begin{aligned}\operatorname{cov}\{m(x)m(x')\} &= kI_q[\min(x, x'), s - \min(x, x') + 1] \\ &\quad - kI_q(x, s - x + 1)I_q(x', s - x + 1) \\ &\quad + ksI_q(x, s - x + 1)I_q(x', s - x + 1) \\ &\quad - ks(1 - q)I_q(x, s - x)I_q(x', s - x') \\ &\quad - ksqI_q(x - 1, s - x + 1)I_q(x' - 1, s - x' + 1).\end{aligned}$$

Derivation of this result is given in section 3. Applications to some practical problems are discussed in section 5. Some comments and computations relating to the usefulness of the asymptotic results are given in section 4.

3. Derivation of Result

For simplicity, we suppose $D = kqs$. Clearly this is permissible for N large. Consider a population of ks items kqs of which $(0 \leq q \leq 1)$ are defective (ks and kqs integral). Suppose we allocate the ks items to k groups, each of size s, at random. Suppose we define a defective group as one in which there are at least x defective items $(1 \leq x \leq s)$. Now we ask for the probability distribution of m, the number of defective groups; we use m to denote both the random variable and some arbitrary one of the values the random variable can assume.

It is obvious that

$$m \geq \max\left\{\left[\frac{D - k(x - 1) + (s - x)}{s - x + 1}\right], 0\right\}.$$

It is also obvious that

$$m \leq \min\left\{k, \left[\frac{kqs}{x}\right]\right\}.$$

Suppose we identify the k groups as 1st, 2nd, 3rd, $\cdots$, ith, $\cdots$, kth and denote the number of defectives in the ith group by d_i.

Then from the above, the probability of $d_1, d_2, \cdots, d_k$ is given by

$$p_{kqs}(d_1, d_2, \cdots, d_k) = \left[\prod_{i=1}^{k}\binom{s}{d_i}\right] \Big/ \binom{ks}{kqs} \tag{3.1}$$

where we require $\sum_1^k d_i = kqs$.

We get $p\{m(x)\}$, the probability that m groups are defective, if at least x out of the s items in a group being defective defines a defective group, by writing the probability that the 1st m groups be defective groups and the last $k - m$ non-defective groups and then multiplying this probability by $\binom{k}{m}$

Thus

$$p_{kqs}\{m(x)\} = \frac{\binom{k}{m}}{\binom{ks}{kqs}} \sum \prod_{i=1}^{m} \binom{s}{d_i} \prod_{i=m+1}^{k} \binom{s}{d_i} \tag{3.2}$$

$$\sum_1^k d_i = kqs$$

$$d_i \geq x, \qquad i = 1, 2, \cdots, m$$

$$d_i < x, \qquad i = m+1, \cdots, k$$

Considerable attention was given the case $x = 1$, which is computationally simplest. Even in this relatively simple case we were unable to show the asymptotic normality of m directly from $p_{kqs}\{m(1)\}$ except for the special case $s = 2$.

To find the asymptotic distribution of m for any x or the joint distribution of any set of the $m(x)$, we proceed indirectly.

Define r_i as the number of lots (of size s) out of k that contain exactly "i" defectives, $i = 0, 1, 2, \cdots, s$. It is obvious that

$$p_{kqs}(r_0, r_1, \cdots, r_s) = \left[\frac{k!}{\prod_{i=0}^{s} r_i!} \prod_{i=0}^{s} \binom{s}{i}^{r_i}\right] \Big/ \binom{ks}{kqs} \tag{3.3}$$

and that we must have

$$r_0 + r_1 + \cdots + r_s = k$$

$$r_1 + 2r_2 + \cdots + sr_s = kqs.$$

We note that the joint distribution given above is almost identical with one given by Freund and Posner [1] for slightly different occupancy problem.

To proceed further, we need to know the proper norming for the r_i, which will presumably involve the means and standard deviations of the r_i.

To this end, define

$$X_{ij} = 1, \quad d_j = i$$

$$= 0, \quad \text{otherwise}, \qquad j = 1, 2, \cdots, k, \tag{3.4}$$

$$i = 0, 1, 2, \cdots, s.$$

Thus

$$E\{r_i\} = E \sum_{j=1}^{k} X_{ij} = kEX_{i1}$$

$$= k \Pr\{d_1 = i\}. \tag{3.5}$$

It is apparent that

$$\Pr\{d_1 = i\} = \binom{ks - kqs}{s - i}\binom{kqs}{i} \Big/ \binom{ks}{s} \tag{3.6}$$

and we find, excluding terms of $0(1/k)$ or less, that

$$k \Pr \{d_1 = i\} \cong k\binom{s}{i}q^i(1-q)^{s-i}. \tag{3.7}$$

Using

$$\begin{aligned} E\{r_i^2\} &= E\Big(\sum_{j=1}^{k} X_{ij}\Big)^2 \\ &= kEX_{i1} + k(k-1)X_{i1}X_{i2} \\ &= k \Pr \{d_1 = i\} + k(k-1) \Pr \{d_1 = i, d_2 = i\}, \end{aligned}$$

it is easy to verify that Var r_i is $0(k)$; we omit the exact expression since it is of no intrinsic interest and not important to the asymptotic argument.

We are thus led to define

$$z_i = \frac{r_i - k\binom{s}{i}q^i(1-q)^{s-i}}{\sqrt{k}}. \tag{3.8}$$

Because of the constraints on the r_i it is clear that a non-singular asymptotic distribution exists only for at most $s-1$ of the z_i, say $z_1, \cdots, z_{s-1}$. Making the indicated transformations and going through a routine argument using Stirling's approximation, we get

$$\begin{aligned} p(z_1, \cdots, z_{s-1}) = {} & \frac{\sqrt{s}}{(2\pi)^{(s-1)/2}(1-q)^{(s-1)/2}q^{(s-1)/2}} \\ & \cdot \exp -\frac{1}{2}\left\{\sum_{i=1}^{s-1}\frac{z_i^2}{f_i} + \frac{\Big[\sum_{i=1}^{s-1}\frac{(s-1)}{s}z_i\Big]^2}{(1-q)^s} + \frac{\Big[\sum_1^{s-1}\frac{i}{s}z_i\Big]^2}{q^s}\right\} \end{aligned} \tag{3.9}$$

where

$$f_i = \binom{s}{i}q^i(1-q)^{s-i}.$$

To obtain the joint distribution of the $m(x)$ we note that

$$\begin{aligned} m(x) &= \sum_{i=x}^{s} r_i = \sum_{i=x}^{s-1} r_i + kq - \sum_1^{s-1}\frac{i}{s}r_i \\ &= \sum_{i=x}^{s-1}\frac{(s-i)}{s}r_i - \sum_1^{x-1}\frac{i}{s}r_i + kq \quad \text{if} \quad 2 \le x \le s-1, \\ &= kq - \sum_1^{s-1}\frac{i}{s}r_i, \quad \text{if} \quad x = s, \\ &= \sum_{i=1}^{s-1}\frac{(s-i)}{s}r_i + kq, \quad \text{if} \quad x = 1, \\ &\equiv k, \quad \text{if} \quad x = 0. \end{aligned} \tag{3.10}$$

Thus any $m(x)$ $(x \geq 1)$ is simply a linear combination of the z_i and as such is asymptotically normal; similarly the joint asymptotic distribution of any $(s - 1)$ of the $m(x)$ is multivariate normal. It can be verified that the variances and covariances of the $m(x)$ are as given in the summary, both by the type of computation indicated to get the variance of the r_i , and by inversion of the covariance matrix of the distribution (3.9) and computation using (3.10).

For the applications we have in mind, a confidence bound on q, the proportion of defective items, must provide a confidence bound on the cumulative probability function of $m(x)$. This will be possible providing the cumulative of $m(x)$ is a monotone function of q (i.e. D). It is intuitively obvious that the cumulative of $m(x)$ is a monotone decreasing function of D, but it has not been possible to show this analytically. In lieu of a proof, Table 1 below is given which demonstrates the monotone relationship for the special case $s = 4$, $k = 12$, $x = 3$.

TABLE 1

Prob $\{m(x) \leq m_0 \mid D, s, k\}$ *for Selected Values of* m_0 *and* D.
$s = 4, k = 12, x = 3$

	m_0		
D	3	6	8
0	1.00000	1.00000	1.00000
13	.99998	1.00000	1.00000
14	.99986	1.00000	1.00000
24	.43659	.99940	1.00000
29	.00346	.80402	.99952
31	.00000	.44398	.97961
33	.00000	.10095	.86508
35	.00000	.00626	.51159
41	.00000	.00000	.00000
48	.00000	.00000	.00000

4. Adequacy of the Asymptotic Theory for Small k

A study of the adequacy of the asymptotic theory for small k is of course complicated by the computational complexity of the exact distribution of $m(x)$. There are additional complications in that both the skewness and the range of the distribution of $m(x)$ depend not only on D but also on x, s and k. To illustrate these points, we give in Table 2 a numerical comparison of the adequency of the normal approximation for cases that differ only in the definition of a defective group, i.e. have different x values. Asymptotic mean values and variances were used in this computation. Agreement of exact and approximate results would presumably be improved by use of exact moments. An empirical continuity correction of .5 has been added to each value of m. Probabilities quoted should be correct to four decimal places.

In Table 2 we note that the distributions of $m(1)$ and $m(4)$ are markedly more skew and have shorter ranges than the distributions of $m(2)$ and $m(3)$. In no case is the normal approximation too bad. Providing one is not concerned

TABLE 2

Comparison of Asymptotic Approximation To Exact Cumulative Distribution of $m(x)$
For $x = 1, 2, 3$ and 4
$s = 4$, $D = 24$, $k = 12$

Pr $\{m(1) \leq i\}$			Pr $\{m(2) \leq i\}$			Pr $\{m(3) \leq i\}$			Pr $\{m(4) \leq i\}$		
i	*Exact*	*Approx.*	i	*Exact*	*Approx.*	i	*Exact*	*Approx.*	i	*Exact*	*Approx.*
						0	.0001	.0000	0	.4709	.3639
						1	.0076	.0009	1	.8880	.8517
						2	.1042	.0409	2	.9919	.9926
						3	.4366	.3632	3	.9998	.9999
			4	.0000	.0000	4	.8221	.8519	4	1.0000	1.0000
			5	.0006	.0018	5	.9791	.9927	5	1.0000	1.0000
6	.0000	.0000	6	.0209	.0318	6	.9994	.9999	6	1.0000	1.0000
7	.0000	.0000	7	.1779	.2136	7	1.0000	1.0000			
8	.0002	.0001	8	.5634	.6041	8	1.0000	1.0000			
9	.0081	.0074	9	.8958	.9072						
10	.1120	.1483	10	.9865	.9914						
11	.5351	.6361	11	.9999	.9997						
12	1.0000	.9591	12	1.0000	1.0000						

with very high or very low probabilities, a practical working rule is to use the normal approximation only if the extremes of possible values of $m(x)$ are at least three standard deviations from the mean. This rule would exclude the use of the approximation for the distributions of $m(1)$ or $m(4)$ in Table 2. The rule appears to be conservative.

5. Statistical Applications of the Distribution of $m(x)$

The random assembly problem we have studied arises naturally and frequently and the various engineering methods used to provide assurance against too many defective lots are not unlike the method proposed here. For example, in designing hold down studs for reactor control drive mechanisms, an engineer designs each stud to withstand the maximum stress it is expected the stud will experience in use. However, he incorporates a safety margin by designing so that two, say, instead of four studs will contain the mechanism at the maximum stress. Consciously, if perhaps intuitively and imprecisely, he tries to bound the probability of more than two failures in any one of the required sets of 4 studs. Because the number of mechnisms is generally small, the total number of studs required will ordinarily be randomly assembled into sets of 4. Consequently the problem of sample testing, as well as that of designing the studs can be viewed as the problem of providing assurance against too many defective lots by insuring against too many defective studs.

Providing assurance of an arbitrarily low probability of a defective lot instead of assurance of an arbitrarily low proportion of defective lots is not new. It is, in fact, the same kind of protection afforded by variables sampling plans for percent defective. Since any finite lot or finite number of such lots can at best be viewed as a sample from an infinite supply, confidence limits based on continuous distribution theory will relate to parameters of that distribution and therefore only to the probability, or average number, of defective lots. Confidence limits on parameters of the continuous distribution permit such statements as: We are 95% sure that the probability is at least $1 - \alpha$ that the number of defective groups is no greater than n where either α is an arbitrarily low preassigned constant and n is random or $1 - \alpha$ is a random variable with n an arbitrarily low preassigned constant.

In many practical situations, particularly those wherein items are produced and made up into batches sequentially, defectives will not occur randomly but with local positive contagion. That is, if a particular batch is known to contain a defective, the probability of another defective occurring in the same batch is thereby increased. In other words, fewer batches will be defective than would be expected if defectives were randomly distributed [providing the proportion of defective items is at least 0 (k^{-1})]. Also from the very notion of positive contagion we would expect the probabilities for large $m(x)$ to be decreased and the probabilities for small $m(x)$ to be increased; this further suggests that we expect the cumulative of $m(x)$ assuming positive contagion to be always greater than the cumulative assuming randomness. Consequently, we expect assurances based on the random assembly model will be conservative.

Implicit in our discussion is a plan for acceptance-rejection of a lot of batches.

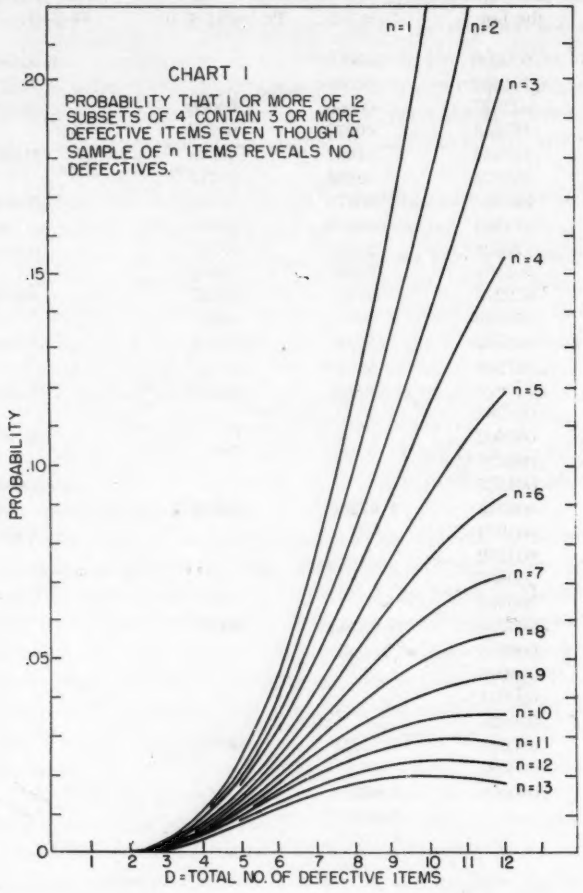
We would like the operating characteristic for such a plan to be dependent solely upon the number of defective *batches* in a lot but we must settle for a dependence upon D, the number of defective *items* in the lot, or some function thereof such as the expected number of defective batches, as in the case of variables sampling, where the quality parameter is average percent defective in a batch.

It is apparent from the above discussion that any attributes sampling plan currently available can be viewed as a plan affording some protection according to any of the criteria mentioned above. We would simply plot the indicated power of a plan against $E\{m(x)\}$, $Pr\{m(x) \leq m_0\}$ for suitable m_0, or any other desired function of overall proportion defective.

TABLE 3

D	Pr of Accepting the Lot	$E\{m(3)\}$	Pr $\{m(3) \leq 0\}$	Pr of Accepting a Lot From an Infinite Process With Fraction Defective $D/48$
0	1.000000	.00000		1.00000
1	.791667	.00000		.81015
2	.623227	.00000	1.00000	.65338
3	.487743	.00278	.99722	
4	.379356	.01092	.98908	.41890
5	.293138	.02683	.97317	
6	.224967	.05273	.94736	.26308
7	.171403	.09066	.90993	
8	.129598	.14246	.83963	.16151
9	.097198	.20981	.79666	
10	.072276	.29417	.72131	.09670
11	.053256	.39082	.63867	
12	.038862	.51921	.54268	.05631
13	.027894	.66142	.44667	
14	.019924	.82498	.35214	.03180
15	.014064			
16	.009802			.01734
17	.006739			
18	.004565			.00909
19	.003043	1.97206	.04042	
20	.001994			.00456
21	.001282			
22	.000807			
23	.000497			
24	.000298	3.65106	.00007	
25	.000174			
26	.000098			
27	.000054			
28	.000028			
29	.000014	5.74635	.00000	
30	.000007			
31	.000000	6.65211		
33		7.57077		
35		8.47641		
41		10.84736		

For a lot size of 48 and an acceptable quality level of 1.5%, Military Standard 105B calls for a normal sample size of 10 with an acceptance number 0 and rejection number 1. The operating characteristic for such a plan is shown in Table 3 below for both individual submitted lot quality and for average submitted lot quality. (This distinction is not always made clear.) Two scalings of the lot number and percent of defectives (for the case where each submitted lot contains D defectives) are shown, namely the expected value of $m(3)$ and $Pr\{m(3) \leq 0\}$. Perhaps a more useful concept of "power" would be the probability of accepting a lot with more than some undesirable number m_0 of sublots containing x or more defectives. This is a perfectly natural way to think of power particularly if the vendor submits lots already packaged into sets (i.e. the random assembly into subsets has already been effected). Hinges are often wrapped in pairs; reactor fuel plates are often submitted in the form of finished fuel assemblies each containing a dozen or more plates. Chart 1 shows for various sample sizes n the probability of accepting a lot of 12 subsets of 4 each when



one or more subset contains 3 or more defectives. The testing is presumed destructive so that the submitted lot size is $48 + n$ although the acceptance-rejection of the lot and the quality thereof refer strictly to the 12 subsets of 4 each. The acceptance number is taken at zero; the rejection number at 1. The chart shows that a sample of 11 is sufficient to keep this probability less than .03. Such a "power function" or "operating characteristic" can be plotted against individual submitted lot quality or against average process quality as desired. It should be noted that in the above example we have considered destructive sampling while earlier discussion was implicitly in terms of non-destructive sampling. It is obvious that asymptotically this is of no consequence; non-asymptotically this difference requires only obvious modifications.

The distribution of $m(x)$ can also be used to decide whether or not random assembly is acceptable as opposed to selective assembly in situations wherein the basic unit of interest is a sublot of items rather than an individual item.

First, however, it would be necessary to decide upon a basis for acceptability of random assembly. We might, for example, use random assembly if $E\{m(x)\} \leq m_0$. In such case, since $E\{m(x)\}$ is obviously monotone increasing in q, the solution of $E\{m(x)\} = m_0$ for q would give an upper bound to allowable grand lot proportion defective warranting random assembly. Alternatively, one might use random assembly if

$$\Pr\{m(x) \leq m_0\} \geq 1 - \alpha$$

for suitably small m_0 and α. Again an upper limit to q be given (approximately) by solving

$$\Pr\left\{\frac{m(x) - E\{m(x)\}}{\sqrt{\text{Var}\{m(x)\}}} \leq \frac{m_0 - E\{m(x)\}}{\sqrt{\text{Var}\{m(x)\}}}\right\} = 1 - \alpha$$

using the normal approximation.

Reference

[1] Freund, J. and Posner, A., "Some Results on Restricted Occupancy Theory", *Annals of Mathematical Statistics*, Vol. 27 (1956), 537–540.

VOL. 3, NO. 1 TECHNOMETRICS FEBRUARY, 1961

Outliers in Patterned Experiments: A Strategic Appraisal

IRWIN D. J. BROSS

Roswell Park Memorial Institute

Two papers "Rejection of Outliers" by F. J. Anscombe and "Locating Outliers in Factorial Experiments" by C. Daniel were presented at the Christmas, 1959 meetings of the American Statistical Association. These papers, plus a discussion by William Kruskal, Thomas S. Ferguson, John W. Tukey and E. J. Gumbel were published in the May 1960 issue of *Technometrics*. The following paper by Dr. Bross continues the exchange of ideas on the subject of "Outliers".

I. THE NEED FOR A STRATEGIC APPRAISAL

Statisticians are in the business of giving advice. One of their favorite pieces of advice to research workers is to carefully plan their studies before undertaking them. What are the objectives of the study? What are the salient features of the experimental situation that will raise obstacles to the attainment of these objectives? How can these obstacles be avoided or overcome? In other words statisticians usually advise their customers to make thoughtful strategic appraisal before getting down to work. This is good advice. It's too bad statisticians do not follow this advice more often in their own work!

What I should like to do here is to make a strategic appraisal of a statistical problem—the problem of outliers in patterned experiments. The first extensive discussion of this problem appeared in the May, 1960, issue of *Technometrics*—though the problem has been troubling investigators for at least 40 years. While the *Technometrics* discussion represents a first step towards what Tukey (1) calls a "return to the elementary problems of statistics", I regard the papers by Anscombe (2) and Daniel (3) as something of a false start. Unless a strategic reappraisal is made, the return to elementary problems will yield only an elaboration of old solutions rather than the new solutions that are needed.

II. WHAT IS AN OUTLIER?

One of the main dangers in making a strategic appraisal is that it is necessary to deal with abstract concepts and in doing so we may drift away from reality. To avoid this, it is a good idea to have a concrete example at hand. For this purpose I shall use the artificial "data" given by Daniel (3, Table I). Although Daniel did not do so, I will suppose that in the array of Table I the rows represent levels of chemical A, the columns represent levels of chemical B, and the entries represent the yield of some product. I have also rearranged Daniel's table slightly.

TABLE I
Hypothetical Yields

	Levels of A			
Levels of B	35	32	37	40
	29	29	34	36
	25	29	30	20
	19	25	25	35
	22	20	29	29

We might start by asking: In what way does the problem of outliers in patterned experiment such as Table I differ from the original outlier problem in which the 20 observations would be regarded as a random sample from a single population? What are the salient or distinguishing features of the situation in patterned experiments? If the observations represented the random sample we could write them down in any order or sequence. In a patterned experiment however there is a definite array due to the pattern that has been imposed. In other words we start with a set of relationships between the observations—relationships which do not depend on the numerical values but rather on the way in which the experiment was laid out. In addition there is a set of relationships—or pattern—of the numerical values themselves. This pattern is evidently linked to the imposed pattern. We notice in Table I that by and large the numbers increase as we move to the right or upwards.

The above remarks may seen obvious but they have implications which are, perhaps, not quite so obvious. In particular we are forced to reconsider what we mean by an "outlier." In the unpatterned experiment we can concentrate on the numerical value of the outlier—a value which stands off in solitary grandeur. In a patterned experiment the numerical value of the outlier may be similar to that of a number of other observations. In other words we must consider both the numerical value and *where* it occurs. An outlier in a patterned experiment is an observation which, in some sense, disrupts the pattern—a concept which is quite different in *character* from the corresponding concept for the random sample.

Some further implications emerge. Let us try to pin down the definition of an outlier in a patterned experiment. Departure from a pattern presupposes some non-null pattern. We have to *characterize* the pattern before we can define departures from the pattern. In practice we would be limited to fairly simple characterizations though this is a restriction imposed by our existing techniques and the phenomena under study is subject to no such constraints. This brings up another concept which has no counterpart in the original problem—the notion of an *isolated* departure. If an observation were out of line and the *adjacent* observations showed a similar disturbance (though to a lesser degree) we would tend to regard this as a real departure from the presumed pattern (providing the observations were independent). On the other hand if an observation stands out from its near neighbors we would regard it as an outlier.

As in the simple random situation there is the further complication of sampling variation. Due to experimental error we cannot expect to find a perfect pattern in actual data. To qualify as an outlier an observation would have to stick out like a sore thumb—represent a serious break in the pattern. Which leads us to the following working definition of an outlier in a patterned experiment: An outlier is an isolated observation which seriously disrupts the apparent pattern of the data.

III. Attitudes Towards Outliers

Statistical methods are sometimes advertised as the way to avoid the intrusion of personal opinions and prejudices in the interpretation of data. Biostatistical methods do serve this purpose. It would be naive however to suppose that personal attitudes are completely nullified. These attitudes do enter into the choice of statistical methods and into their justification (although this point may be well concealed by the mathematics). Since outliers disrupt the patterns our first impulse is to eliminate them. Our attitude towards outliers cannot help but influence the way in which we set about this process of elimination. Our attitude towards outliers depends in turn upon how we feel about throwing away data in scientific studies. On one hand we may take a "cheerfully detached" attitude and, like the Queen of Hearts, go around shouting "off with their heads!" This attitude is reflected in schemes for truncation of data, elimination of a fixed proportion of the observations, and other procedures which mechanically eliminate the observations. I find this cavalier attitude rather distressing. As Kruskal points out (4) "The immediate pressures of practical statistical analysis are almost uniformly in the direction of suppressing announcement of observations that do not fit the pattern; we must maintain a strong sea-wall against these pressures." In my experience reputable scientists in all fields have an aversion—almost a phobia—against leaving out data. Is the feeling (which I share) merely the result of clinging to old scientific traditions? I don't think so. I think this feeling comes from long experience (some of it recent) with people who are out to prove their point and proceed to do so by eliminating "embarrassing" observations.

I want to make the point that the loss from omitted observations is *not* measured by the increased variance of an individual estimate. So far as publications are concerned, the dropping of observations loses the confidence of a thoughtful reader. Therefore I feel we should be reluctant to drop observations and should do so only for compelling reasons.

We cannot express this reluctance in a formula or as a number, so that it cannot appear directly in the technical manipulations, but I think it is an important factor in the strategic picture. Anscombe's "insurance" formulation omits this factor (2). In the "insurance" formulation the dropping of outliers is a minor premium that is paid to obtain a safeguard against the major damage produced by outliers. Hence the "insurance" view reflects the attitude that outliers are simply a nuisance and the problem is merely to get rid of them without losing too many non-outliers in the process.

IV. Damage Due to Outliers

Now let us turn to the crucial question: What is the principal damage produced by outliers in a patterned experiment?

Most investigators who have used analysis of variance (ANOVA) techniques in patterned experiments have encountered what I will call the "star burst" effect. This effect is especially striking in the factorials which are analyzed down to high order interactions. Sometimes omission of a single observation leads to an ANOVA table which suggests a null or simple non-null situation. However inclusion of this observation gives an ANOVA table that is gaily decorated with stars (* or **) all up and down the lines. An analyst who was unfamiliar with this "star burst" effect would conclude from the uncensored table that all sorts of main effects and interactions are "statistically significant." Hence if the "wild" observation is really meaningless the analysis leads to a complete misconception of the underlying relationships in the data. So we can safely say that an outlier can do a great deal of damage in a patterned experiment.

There is, however, one very important point to note. We may attribute the damage to the outlier but, in fact, it is the *combination* of ANOVA techniques and an outlier that is responsible for the "star burst" effect. If we think of the outlier as a "contaminated" observation then in the original data this contaminator is isolated in a single cell. It is the ANOVA techniques that spread the contamination throughout the results. The contamination gets into the column, row, and grand means and thence is carried into all the estimates and significance tests. Other kinds of analysis (eg. non-parametric) are likely to be much less affected by the presence of the outlier. This suggests that one strategy for dealing with outliers would be to shift to some other type of analysis. However in what follows I will stay with the strategy of eliminating outliers.

It is essential to realize the role of the ANOVA techniques in the outlier problem—that they magnify the damage that the outlier can produce. Moreover the vulnerability of ANOVA techniques is by no means limited to outliers. The techniques are vulnerable to other disturbances or departures from the standard assumption—non-additivity and heterogeneity of variance, for example. Indeed in a given body of data it is quite possible that several different types of disturbances, including outliers, are present. Any device that we set up for eliminating outliers will have to function in the presence of other disturbances and there is a risk of producing unexpected and undesirable side effects.

Turning back to the question of what is the principal damage, there are two conventional answers. The first answer focuses on estimation. The outlier puts a major bias in some of the contrasts and lesser biases in various other contrasts. There is also an increase in the variance of the estimates. It should be noted that the increased variance is not primarily due to the slight reduction in sample size (eg. which would be a multiplicative effect of the order of $n/n-1$) but rather stems from the omission of the most prominent member of the sample. Anscombe has shown that the latter effect is typically 15 to 20 times greater than the former effect. The same remark applies when we turn to the second focus, significance tests. However the effect on individual estimates or tests is of less importance, I think, than the overall damage that is done. The principal

damage is that with standard ANOVA methods the outlier can mess up beyond all recognition the *relationships* between the variables.

It is not easy to assess the overall damage by a simple formula. However I don't think that our desire for a simple mathematical derivation should be allowed to distort our strategic picture of the problem. If disruption of the pattern of relationships is the principal damage, then this should be our focus of attention, even though the tactical problems are difficult.

V Lines of Approach

The above strategic appraisal has, I hope, helped to clarify the nature of the problem of outliers in patterned experiments. If we adopt the strategy of elimination of outliers three questions naturally arise. How can we tell if there are outliers in a given body of data? If there are outliers, how can we identify them? If we eliminate the identified outliers, how should the remaining data be analyzed?

The first question suggests that some type of preliminary analysis for outliers be made. What would be the basis for such analysis? Both Anscombe and Daniel choose as a basis the *residuals*—the departures of the individual observations from their expected values (as calculated by standard ANOVA methods for the given design). In Table I the expected values would be: Expectation = row mean + column mean − grand mean. For example in the lower left hand cell this gives $25 + 26 - 29 = 22$. The observation is also 22 so that the residual is 0. This is a natural basis but let us consider it in the light of the previous strategic appraisal.

First of all we may note that if there is an outlier in the data then *every* residual will be affected by this outlier (Daniel calls them "biased residuals"). Secondly, as Anscombe points out, even without outliers the residuals in a complex design are correlated—sometimes highly correlated. This correlation pattern is not at all obvious to anyone unfamiliar with the details of a specific design. A third point to note is that the calculation of expected values makes tacit assumptions about the pattern—additivity for instance—which can't apply to the outlier and may not apply to the rest of the data. Hence the use of residuals as a basis has some serious tactical and strategic disadvantages. The main advantage is that we are able to stay within the usual ANOVA conceptual framework.

Without going into the technical details, let us see the general route by which we would proceed from the basis. We might consider that the problem has now been reduced to the original outlier problem (with some complications). In other words we could consider only the magnitudes of the deviations and reject those observations whose departures exceed 2σ (or some other multiple of sigma). What we have done, then, is to reduce the outlier problem in a patterned experiment to the outlier problem in a random sample. From a technical view such a reduction is highly desirable. From a stragetic view, however, the key question is: Have we thrown out the baby with the bath? Does this reduction preserve the salient features of the patterned experiment problem? I think not. We have ignored the pattern of the residuals so we have cut out the heart of the problem. This, of course, is my view so it is only fair to add that one of the

discussants (5) considers that this approach will be "the cornerstone on which future generations of outlier rejectors will base their work."

Daniel's approach, while based on residuals, is something more than a reduction to the original problem. He remarks in connection with 2^n factorials "if a defective value which has an expected value near one end of the range of responses actually appears near the other end of the range ··· I see no alternative in such cases to making a series of trials, replacing in turn each of the 2^n responses by the average of the others, and then revising the estimates of all effects accordingly. If one and only one of the resultants sets permits a simple regression equation, containing mostly main effects and interactions among the factors having large effects, then a strong case can be made for naming as defective the value whose revision caused the simplification." While this approach may be too cumbersome for actual applications, it deals directly with the main damage—the disruption of the pattern of relationships produced by an outlier. Also it tacitly takes into account the location of the outlier. It requires the investigator to *look* at his data and exercise a certain amount of judgment rather than to apply an automatic rejection process. As a practical approach it is tactically weak but strategically much stronger than the reduction process.

Having observed some of the difficulties encountered when residuals are used as a basis a natural question is: Do we *have* to use residuals? The outlier problem in patterned experiments represents the breakdown of ANOVA techniques in the presence of a disturbing factor. Would it not be more plausible to turn to techniques outside of the ANOVA conceptual framework to deal with the disturbing factor? Are there feasible alternative approaches? Let's take a look.

VI. An Alternative Line of Attack

Let us review the strategic appraisal to see if it suggests an alternative line of attack. By the working definition, an outlier is an isolated observation that seriously disrupts the apparent pattern in the data. Does this suggest any way to see *if* an outlier is present, to identify the outlier, and to deal with it? If we are looking for a sensible way to proceed rather than a neat mathematical derivation, then I think the definition suggests an alternative line of attack. Let us look back at the data in Table I. The first step is to ask: What is the apparent pattern? Here I think we can say without much difficulty that the observations tend to increase as we go to the right or up in the Table. Technically we might say that there is joint monotonicity. The next step is to ask: Are there any apparent departures from this simple pattern? By scanning the data we become aware that there are several departures, some minor and some serious. For example the lower left hand entry is out of line with the entry above it and to the right of it but it is not very badly out of line in the sense that if it were 18.5 instead of 22 this irregularity would be removed. On the other hand in the center row of the fourth column there is an entry which is distinctly out of line with all of its near neighbors. Moreover to bring this entry back into line it would have to be raised from 20 to 35.5. Here, then, we have an observation which satisfies the definition of an outlier.

The above remarks may call forth two very different responses in readers. The first response, perhaps from persons experienced with data and analysis, would be "This is obvious! Any sensible investigator would do this without all this elaborate strategic appraisal." The second response, perhaps from technically minded readers, would be "But this is absurd! This gives the investigator license to pick and choose the observations he wishes to discard. This procedure is not objective and flatly contradicts the preceeding strategic appraisal."

In answer to the first objection I can only say that the approach does seem to have been "obvious" to the outlier experts. In answer to the second objection I would say that the task is to provide some type of objective criteria based on this line of attack—a tactical problem.

Can the line of attack be formalized? It is not hard to derive a more rigorous formulation (that gets into some unsolved technical problems) but we are thereby led to a different *basis* for rejection.

The central notion in a new formulation is the *pairwise comparison* of the observations in an array. The imposed pattern of the experiment orders the observations—though not necessarily in a *complete* ordering. If we assume some specific pattern then we can speak of the *expected* direction of a pairwise comparison. If the observed direction differs from the expected direction then we call the result an *inversion*. For example, proceeding to the right along the bottom row we expect the numbers to increase. Of the six possible pairwise comparisons in this row, four are in expected direction, one (the entries in the first two columns) is an inversion, and in one case we have a tie. Ties present a technical problem but if they are relatively rare we can simply count them as half an inversion. An inversion count in the table then can be taken as the basis for an outlier elimination procedure.

This formalization is straightforward enough but it tacitly assumes that there is a recognizable pattern in the data. Before proceeding further with this approach we might want some assurance that we are not "seeing things" in the data. There is another reason for checking into the assumed pattern. In a grand null situation or in a situation where a large proportion of the stimulus variables are null, we are getting out of the patterned situation and into the simple random one. Suppose for example that three of the stimulus variables in a 2^n factorial are actually null. Then a given outlier would be one of a set of eight measurements under effectively the same conditions. In such degenerate cases we would think of using a simpler procedure such as replacement of the outlier by the average of the other observations in the set. In what follows I shall suppose that we are in a non-null situation. This point can be readily verified by using procedures which are relatively insensitive to the presence of outliers.

In Table I, for example, we have at our disposal the Sequence Sign Test and other devices that are based on inversions. These procedures are hardly more trouble than simply scanning the data. For example let us consider the rows in Table I. Here an inversion will occur if the right hand member of a paired comparison has a lesser value than the left hand member. Thus in the top row there is one inversion, in the next row there is half of an inversion, in the middle row there are three inversions, in the fourth row half of an inversion and in the

bottom row one and one-half inversions. All told there are $6\frac{1}{2}$ inversions out of $6 \times 5 = 30$ paired comparisons.

To test for a monotone relationship in the rows we subtract the number of inversions, $6\frac{1}{2}$, from the number of non-inversions ($30 - 6\frac{1}{2} = 23\frac{1}{2}$) to get $23\frac{1}{2} - 6\frac{1}{2} = 17$. We subtract one more for continuity correction, giving 16 and square 16 to obtain a numerator of 256. Dividing by the total number of pairwise comparisons, 30, gives $256/30 = 8.53$. Up to this point we have simply carried out an ordinary sign test. However to take account of the fact that the six pairwise comparisons are not independent we must introduce a multiplicative correction factor. The denominator of the correction factor is twice to the number of entries in the rows plus five. Here it would be $2 \times 4 + 5 = 13$. The numerator of the correction factor is 9 so that the value here is 9/13 or 0.69. Multiplying 8.53 by this correction factor gives 5.89. The usual critical values for the sign test apply here so that the 5% level is 3.84 and the 1% level is 6.64. Even with a possible outlier we can reject the hypothesis of a null effect along the rows at the 5% level. A derivation of the Sequence Sign Test can be found in my recent paper on Statistical Criticism (6).

I have dealt with this tactical finesse at some length because it illustrates, for readers not familiar with inversion methods, just how simple and straightforward the tactics can be. Applying the Sequence Sign Test to the columns we find $2\frac{1}{2}$ inversions in the first three columns and 2 inversions in the right

TABLE II

Inversion Tables (For Table I)

TABLE II (a)

Number of Row Inversions by Cell
(in 3 Pairwise Comparisons)

	Levels of A			Total
1	1	0	0	2
$\frac{1}{2}$	$\frac{1}{2}$	0	0	1
1	1	1	3	6
0	$\frac{1}{2}$	$\frac{1}{2}$	0	1
1	1	$\frac{1}{2}$	$\frac{1}{2}$	3

TABLE II (b)

Number of Column Inversions by Cell
(in 4 Pairwise Comparisons)

	Levels of B				
	0	0	0	0	0
	0	$\frac{1}{2}$	0	0	0
	0	$\frac{1}{2}$	0	2	1
	1	0	1	1	1
	1	0	1	1	1
Total	2	1	2	4	

TABLE II (c)

Number of Main Diagonal Inversions by Cell
(Number of Pairwise Comparison in Denominator)

	Levels of A			
Levels of B	—	0/4	0/8	0/12
	0/3	0/5	0/7	0/9
	0/6	0/6	0/6	4.5/6
	0/9	1/7	1/5	0/3
	1/12	0.5/8	1/4	—

TABLE II (d)

Cumulated Inversions by Cell
(Number of Pairwise Comparisons Variable)

		Levels of A				Total
	Levels of B	1	1	0	0	2
		½	1	0	0	1½
		1	1½	1	9½	13
		1	1½	2½	1	6
		3	1½	2½	1½	8½
Total		6½	6½	6	12	31

hand column. The Sequence Sign Test gives a numerical value of 13.5 which is again significant at the 1% level. If the reader takes the trouble to go through the actual calculations he will be struck, I think, that by the fact that pairwise comparisons involving one of the cells account for a sizable proportion of all the inversions. This point can be made clearer by counting in a slightly different fashion. If we count each inversion twice, once for each member of the pair, we obtain the inversion tables shown in Table II. The point of this symmetrical method of counting inversions is that we can associate a count with each cell in the original table. Table II(a) shows the inversions for the rows, Table II(b) for the column, and Table II(c) for the main diagonal comparisons which we will come to shortly. Table II (d) merely sums the inversions in the three previous tables. Inspection of Table II suggests very strongly that there is an outlier in this data. Indeed it locates the trouble spot for us. We see that the central row and the right hand column give the most inversions. So we would suspect that an outlier occurs in the middle row of the right hand column. All that has really been done here is to systematize the scanning approach which was introduced at the start of this section. If we want formal, objective elimination rules an inversion count in the form of Table II provides a basis.

Table II (c) concerns main diagonal comparisons, that is comparison of an observation with the observations below and to the left of it in the table and the set of observations above and to the right of it in the table. It should be noted that joint monotonicity specifies the direction of these comparisons but

it does not specify the direction of the minor diagonal comparisons. Minor diagonal comparisons provide information on the relative strength of the two factors—a question I don't wish to get into here. They are therefore omitted. It is interesting that in a table of main diagonal inversions the *only* inversions that occur involve the suspected outlier. In general if a simple joint monotone relationship exists we would expect to find the proportion of inversions going down as the pairs in the comparison are farther apart. This fact might be used to distinguish outliers from minor discrepancies in the pattern.

VII. Tactical Development

While I will not go into the tactical development, I do want to say a little about the general lines that might be taken. One approach would be to focus on the proportion of all inversions that were produced by a given observation. For example we could divide the number of inversions by the number of cells. If a given observation produces substantially more inversions than its "quota" it would be a candidate for elimination. For example we have 31 inversions in all in Table II (d) and the quota would be $31.0/20 = 1.55$. Our suspect thus exceeds its quota many times over $(9/1.55 = 5.81)$.

At this point it might be objected that the inversion approach does not take into account the actual magnitude of the observation. This is a valid objection but I don't feel that we are under any compulsion to confine ourselves to a single approach—even though this would simplify the theory. We can bring in the magnitude by considering how much the observation would have to be changed in order to eliminate the inversion. If this change was of the order of magnitude of several standard errors this would provide an additional reason for rejection. Since I take a reluctant attitude towards rejection of data I would be in favour of basing the rejection on all the information that would be relevant to the decision. To be sure this presents difficulties in the calculation of rejection rates and other performance measures but I regard this as a secondary consideration.

To take direct account of the numerical values the procedure that Tukey calls "Winsorization" (1) can be used. If we eliminate the middle entry in the right hand column of Table I what value would we use in its place to avoid the inversions? By using a number greater than 30 we would avoid the row inversions. By using a number greater than 35 and less than 36 we could avoid the column inversions. By using a number greater than 29 we could avoid the diagonal inversions. In general we would obtain in this way a set of inequalities (a set depending on the number of variables and the location of the outlier). To obtain a definite replacement we could use, say, the average of the upper and lower limits. In this example this would be the average of 35 and 36 or 35.5.

To avoid inversions we have to raise the entry from 20 to 35.5—or 15.5 units. If we had previous experience with the experimental error in similar experiments we could compare the change, 15.5, with the previous estimate of σ. Lacking past experience we could calculate the error mean square for the patterned experiment (using the replacement if needs be). Daniel gives the mean square as 2.47 (and the square root would be 1.57) so we see that the change is 10 times

the standard error. We can build up a strong case for elimination of the outlier by utilizing additional information in this way.

The process of Winsorization also provides a plausible answer to the question: If we eliminate the outlier, how should the data be analyzed? If we want to use ANOVA techniques we can restore the balance by replacing the outlier by its Winsorized estimate. Doubtless this suggestion will not sit well with the devotees of current missing plot techniques but we have already swallowed a camel so why strain at a gnat? I would like to stress that it is the use of ANOVA techniques that makes it important to replace the outlier with some fictitious value. The techniques related to the Sequence Sign Test can be applied by merely omitting the outlier.

Significance tests and formal rules for rejection of outliers could be based on inversions and this area is open for tactical development. There is the technical difficulty that the theoretic distribution of inversions for non-null cases is not known. One possibility would be to apply the usual outlier theory to the empirical distribution of inversions (eg. Table II) but there are two technical hitches. As noted earlier, in the non-null case the probability of an inversion would depend on the location of the two entries in the pairwise comparison. Also the entries in Table II are not independent. There is the further problem of combining row, column, and main diagonal inversions. While it would be nice to have a formal test here I have the feeling that when we have a bad outlier (as here) then a formal test is really gilding the lily.

If we are willing to discard some of the information it is not hard to set up an approximate significance test. Leaving out the row of the suspected outlier we have (counting inversions in the usual, non symmetrical manner) 1, $\frac{1}{2}$, $\frac{1}{2}$, and $1\frac{1}{2}$ row inversions. The mean of this series is 7/8 and the usual estimate of variance would be 11/48. Leaving out the column of the suspected outlier we have 1, $\frac{1}{2}$, and 1 column inversions. The mean of this series is 5/6 and the variance 1/12. Now consider a total of the inversions in a row and those in a column. Call this total Z. We would then expect Z to be equal to $7/8 + 5/6 = 1.71$. We would estimate the variance of Z as $11/48 + 1/12 = .31$ and the standard error as .56. So by the usual 2σ rule we would expect Z to lie between 0 and 2.83. The sum of the row and column inversions for the row and column containing the suspected outlier is $3 + 2 = 5$ which lies well outside the expected range.

I don't want to give the impression that tactical development along these lines is either easy or unnecessary. I have mentioned the nasty problem of partially ordered patterns. In the example I have assumed a natural ordering (by amount of chemical) but in many patterned experiments there may be only partial ordering or even no ordering at all. In this case we can use the response variable for the orderings but this gets us into various complications. A further difficulty comes up in patterned experiments where simple monotonic relationships are unlikely. I think that the nature of the patterns likely to be found in a given class of data is a determining factor in the outlier methodology we would want to use. Finally I want to mention again the problem of simultaneous occurrence of other disturbances apart from outliers. While the inversion approach is probably less affected by such disturbances than an approach based on residuals, it is not completely immune.

VIII. Summary

A strategic analysis of the problem of outliers in patterned experiments reveals serious shortcomings in the current approaches and suggests a rather different line of attack. While this strategic appraisal is open to arguments and objections—as almost any such appraisal will be—I believe it will serve to illustrate the need for such appraisals, the nature of such appraisals, and the practical value in making a strategic survey before undertaking tactical development. In this case the outlines of a tactical development emerge from the strategic appraisal—I must confess in a way not expected when I started the appraisal. The strategic appraisal is defective in the sense that it does not follow through with the tactical details—and a strategic appraisal is of limited practical value until there is a tactical follow-up. Many of the current papers in the statistical literature are very strong tactically but are strategically unsound so a paper with the opposite fault at least contributes some variety.

IX. References

1) Tukey, J. W., (1960), Discussion of the Papers of Messrs. Anscombe and Daniel. *Technometrics*, 6, 160.
2) Anscombe, F. J., (1960), Rejection of Outliers. *Technometrics*, 6, 123.
3) Daniel, C., (1960), Locating Outliers in Factorial Experiments. *Technometrics*, 6, 149.
4) Kruskal, W. H., (1960), Discussion of the Papers of Messrs. Anscombe and Daniel. *Technometrics*, 6, 157.
5) Ferguson, T. S., (1960), Discussion of the Papers of Messrs. Anscombe and Daniel. *Technometrics*, 6, 159.
6) Bross, I. D. J., (1960), Statistical Criticism. *Cancer*, 13, 394–400.

Multiple Comparisons with a Control in Balanced Incomplete Block Designs

D. S. ROBSON
Cornell University

This paper shows that Dunnett's procedure for making multiple comparisons of several treatments against a single control can be used on the adjusted treatment means from balanced incomplete block designs.

Dunnett (1955) published tables for a one- or two-sided multiple comparison of p treatments with a control. His tables are constructed for the case where the p treatment means $\bar{X}_1, \cdots, \bar{X}_p$ and the control mean $\bar{X}_0$ are mutually independent and normally distributed with common variance σ^2/K estimated by s^2/K, where K is a known constant and s^2/σ^2 is chi-square distributed with v degrees of freedom, independent of the $p + 1$ means. Under these assumptions the chance variables

$$Z_i = \frac{(\bar{X}_i - \bar{X}_0) - (\mu_i - \mu_0)}{\sqrt{2/K}}$$

for $i = 1, \cdots, p$, have the multivariate normal distribution with means 0 and variances σ^2 and with a correlation of $\frac{1}{2}$ between Z_i and Z_j. Utilizing tables prepared by the National Bureau of Standards, Dunnett then tabulated the values of t_1 and t_2 for which

$$P_1 = \Pr(Z_1 < t_1 s, Z_2 < t_1 s, \cdots, Z_p < t_1 s)$$

$$P_2 = \Pr(|Z_1| < t_2 s, |Z_2| < t_2 s, \cdots, |Z_p| < t_2 s)$$

for $P_1, P_2 = .95$ and $.99, p = 1, 2, \cdots, 9$, and $v = 5(1)20, 24, 30, 40, 60, 120, \infty$.

The application of Dunnett's test to the adjusted treatment means from a balanced incomplete block design (utilizing only the intrablock information) might at first appear questionable since in this case the $p + 1$ adjusted means $\bar{X}_0', \bar{X}_1', \cdots, \bar{X}_p'$ are not independent. Because of the balanced condition, however, all adjusted treatment means have the same variance and all pairs have the same covariance. Consequently, for $0 < i < j$

$$\begin{aligned}\sigma^2(\bar{X}_i' - \bar{X}_0') &= 2[\sigma^2(\bar{X}_0') - \sigma(\bar{X}_i', \bar{X}_0')] \\ &= 2\sigma(\bar{X}_i' - \bar{X}_0', \bar{X}_j' - \bar{X}_0') \\ &= 2\sigma^2/rE\end{aligned}$$

where σ^2 is the intrablock error variance, r is the number of replicates on each treatment and E is the efficiency factor of the design. The adjusted Z variables

$$Z_i' = \frac{(\bar{X}_i' - \bar{X}_0') - (\mu_i - \mu_0)}{\sqrt{2/rE}}$$

are therefore also normally distributed with means 0, variances σ^2 and pairwise correlations of $\frac{1}{2}$.

This justification for the use of Dunnett's test with balanced incomplete blick designs also implies the applicability of other procedures of this type; e.g., those of Tukey (1953), Bechhofer (1954), and Duncan (1955).

Numerical Example

An application of Dunnett's test to a doubly balanced incomplete block design is illustrated below for a taste-testing experiment conducted by D. H. Wallace of Cornell University. The doubly balanced design in which each pair of treatments occurs together in an equal number λ of blocks and each triplet is together in δ blocks was constructed by Calvin (1954) to eliminate from error the correlation effect common to the ith and jth treatment when both are in the same block. The present experiment, in which $p = 9$ cabbage varieties were to be compared with a standard, was designed with the parameters $\lambda = 4$, $\delta = 1$ and

$q = 30$ = number of blocks (tasters)

$k = 4$ = number of varieties per block

$r = 12$ = number of replicates

giving

$$E = \frac{\lambda(p+1)}{rk} = 5/6$$

Raw cabbage samples presented to the tasters were scored for texture on a five point scale.

Table 1

Analysis of variance of cabbage texture scores.

Scorce	d.f.	S.S.	M.S.	F
Total	119	135.325		
Blocks (unadj.)	29	16.575	.5716	
Varieties (adj.)	9	48.127	5.3474	6.82**
Correlations (adj.)	36	35.333	.9815	1.25
Error	45	35.290	.7842	

The experimenter in this case wished to determine which cabbage varieties, if any, differed from the standard variety in average texture score. Adjusted treatment means were computed according to the procedure outlined by Calvin and the deviations $\bar{X}_i' - \bar{X}_0'$ from the control were tabulated in rank order, as shown in Table 2. Dunnett's tables were entered to find the value of t_2 at the 95% level, which for $p = 9$ treatments and 45 degrees of freedom is $t_2 = 2.90$, and the tabulated differences were then compared with the allowed difference of

$$A = \pm t_2\sqrt{2s^2/rE}$$
$$= \pm 2.90\sqrt{2(.7842)/10}$$
$$= \pm 1.15$$

TABLE 2

Ranked deviations of adjusted treatment means from the adjusted control mean

i	1	2	3	4	5	6	7	8	9
$\bar{X}'_i - \bar{X}'_6$	−1.43	−.60	−.28	−.20	−.08	.15	.68	.75	1.10

Only the lowest of the 9 treatment means was significantly different from the control mean.

REFERENCES

BECHHOFER, R. E. (1954), "A single-sample multiple decision procedure for ranking means of normal populations with known variances," *Ann. Math. Statist.*, 25, 16–39.

DUNCAN, D. B. (1955), "Multiple range and multiple *F* tests," *Biometrics*, 11, 1–42.

DUNNETT, C. (1955), "A multiple comparisons procedure for comparing several treatments with a control," *J. Am. Stat. Assoc.*, 50, 1096–1121.

TUKEY, J. W. (1953), "The problem of multiple comparisons," Unpublished note, Princeton University, Princeton, N. J.

Estimates of Bounded Relative Error for the Mean Life of an Exponential Distribution[1]

BENJAMIN EPSTEIN

Wayne State University and Stanford University

1. THE PROBLEM

Consider a random variable X having p.d.f.

$$f(x; \theta) = \frac{1}{\theta} e^{-x/\theta}, \qquad x > 0, \qquad \theta > 0. \tag{1}$$

In this note we find an estimation procedure for the mean life θ having a small relative error. Put more precisely, we show how to get an estimate which is, with some preassigned confidence $1 - \alpha$, within a certain percentage (100 δ percent) of the true but unknown mean life θ. In the language of life testing, this will require observing a suitably large number of failures, r.

The exact solution of the problem involves considerations like those in [1]. Suppose that a fixed number, r, of failures is observed, and suppose that the associated total life is T_r, then it can be shown that the "best" estimator of θ in the minimax sense corresponding to the loss function

$$\begin{aligned} L(\theta, a) &= 0, \quad \text{if} \quad 1 - \delta \leq \frac{a}{\theta} \leq 1 + \delta \\ &= 1, \quad \text{otherwise} \end{aligned} \tag{2}$$

is given by the estimator

$$a = 2\delta \, T_r/r \log\left(\frac{1 + \delta}{1 - \delta}\right). \tag{3}$$

However, for $0 < \delta \leq .1$, $a \sim \hat{\theta}_r = T_r/r$, since

$$a \sim \frac{T_r}{r}\left(1 - \frac{\delta^2}{3} + o(\delta^2)\right). \tag{4}$$

Equation (4) suggests that when $\delta \leq .1$ and $\alpha \leq .1$, it is reasonable to ask for what values of r is

$$\Pr\left(\left|\frac{\hat{\theta}_r - \theta}{\theta}\right| \leq \delta\right) \geq 1 - \alpha? \tag{5}$$

Since the r required to satisfy (5) is large when $\delta \leq .1$ and $\alpha \leq .1$, we may safely assume that $\sqrt{r}\,(\hat{\theta}_r - \theta)/\theta$ is approximately distributed as $N(0, 1)$.

---

[1] Work done with partial support from the Office of Naval Research.

Thus to meet the conditions imposed by (5) means that r must be chosen in such a way that

$$r \geq c_\alpha^2/\delta^2 \tag{6}$$

where

$$\begin{aligned} c_\alpha &= 2.576 \quad \text{if} \quad \alpha = .01 \\ &= 1.960 \quad \text{if} \quad \alpha = .05 \\ &= 1.645 \quad \text{if} \quad \alpha = .10. \end{aligned}$$

If $\delta = .01, .05, .10$ and $\alpha = .01, .05, .10$ the values of r required are tabulated in Table 1.

TABLE 1

Values of r such that $\Pr\left(\left|\frac{\hat{\theta}_r - \theta}{\theta}\right| \leq \delta\right) = 1 - \alpha$.

	α		
δ	.01	.05	.10
.01	66,400	38,400	27,100
.05	2654	1537	1082
.10	664	384	271

It is of interest to show how one can compute, for any preassigned r, the confidence level of the assertion that $1 - \delta \leq a/\theta \leq 1 + \delta$ where $a = 2\,\delta T_r/r \log (1 + \delta/1 - \delta)$. This is of particular importance when r is small (and hence asymptotic theory is no longer valid). It can be verified that for any preassigned r

$$\begin{aligned} \Pr\left(1 - \delta \leq \frac{a}{\theta} \leq 1 + \delta\right) &= \int_{c_1(r,\delta)}^{c_2(r,\delta)} \frac{x^{r-1}e^{-x}}{r!}\,dx \\ &= \sum_{k=r}^{\infty} p(k; c_2(r, \delta)) - \sum_{k=r}^{\infty} p(k; c_1(r, \delta)) \end{aligned} \tag{7}$$

where

$$\begin{aligned} c_1(r, \delta) &= (1 - \delta)\,\frac{r}{2\delta} \log\left(\frac{1 + \delta}{1 - \delta}\right) \\ c_2(r, \delta) &= (1 + \delta)\,\frac{r}{2\delta} \log\left(\frac{1 + \delta}{1 - \delta}\right). \end{aligned} \tag{8}$$

2. Numerical Examples

1. How many tubes should be tested in order that there is a probability of at least .90 that the estimate is within 10% of the true mean life?

Solution: In the notation of (5) and (6), $\delta = .1$, $\alpha = .1$, and $c_\alpha = 1.645$. Therefore the number of tubes tested should be $\geq (100)(1.645)^2 = 271$. If

the underlying distribution is exponential this means that we must observe at least 271 failures in order to get an estimate such that we can be 90% confident that the estimator is within 10% of the true but unknown mean life.

2. We have available information from a life test in which 5 failures occurred with associated total life $T_r = 1000$. Assuming an exponential distribution, find the minimax estimator a associated with the loss function

$$L(\theta, a) = 0 \quad \text{if} \quad .8 \leq \frac{a}{\theta} \leq 1.2$$

$$= 1, \quad \text{otherwise.}$$

Also compute the confidence that we will have in the correctness of the assertion that $.8 \leq a/\theta \leq 1.2$.

Solution: From (3), the minimax estimator of θ based on the 5 failures is given by $a = 2(.2)(200)/\log(1.2/.8) = 80/\log(1.5) = 197$. To find the confidence in our assertion that $.8 \leq a/\theta \leq 1.2$, we use (7) and (8). This gives

$$\text{Confidence} = \sum_{k=5}^{\infty} p(k; 6.08) - \sum_{k=5}^{\infty} p(k; 4.06) = .7255 - .3829 = .3426.$$

Reference

[1] M. A. Girshick, H. Rubin, and R. Sitgreaves, "Estimates of Bounded Relative Error in Particle Counting," *Annals of Mathematical Statistics 26*, 276–285, 1955.

On Methods of Constructing Sets of Mutually Orthogonal Latin Squares Using a Computer. II*

R. C. BOSE, I. M. CHAKRAVARTI and D. E. KNUTH

University of North Carolina and Case Institute of Technology

A new class of triplets of mutually orthogonal latin squares is derived. These are the consequence of attempts to extend the results described in Part I of this paper which appears in *Technometrics*, 1960, Vol. 2, No. 4, p. 507.

1. *Introduction, Notations and Motivation*

In [1] a method of searching for sets of mutually orthogonal Latin squares of size $v = 4t$ where $4t$ is the order of a Hadamard matrix is given. The method is based on the concept of orthogonal mappings of a group, due to Mann [4]. For the group G of order $4t$, the module $G(2, 2t)$ whose elements are vectors (a, b) where a is a residue class (mod 2) and b is a residue class (mod $2t$) is used. The group operation (addition) is defined by

$$(a_1 , b_1) + (a_2 , b_2) = (c_1 , c_2)$$

where $a_1 + a_2 = c_1 \pmod 2$ and $b_1 + b_2 = c_2 \pmod{2t}$. The standard order of the elements $x_i = (a_i , b_i)$ of $G(2, 2t)$ was defined to be

$$(0, 0), (0, 1), \cdots , (0, 2t - 1), (1, 0), \cdots , (1, 2t - 1).$$

We also use the notation

$$P_1(x_i) = a_i , \qquad P_2(x_i) = b_i$$

and then we have

$$\begin{aligned} P_1(x_i) = a_i &= 0 && \pmod 2 \text{ if } \quad 1 \leq j \leq 2t \\ &= 1 && \pmod 2 \text{ if } \quad 2t + 1 \leq j \leq 4t \\ P_2(x_i) = b_i &= (j - 1) && \pmod{2t} \text{ if } \quad 1 \leq j \leq 2t \\ &= j - 1 - 2t && \pmod{2t} \text{ if } \quad 2t + 1 \leq j \leq 4t \end{aligned}$$

The $4t$ rows of a square are made to correspond (1, 1) to the $4t$ elements $x_1 , x_2 , \cdots , x_{4t}$ of G and also the $4t$ columns. The cell belonging to the row x_i and column x_j is said to be the cell (x_i , x_j) and the basic square I has the entry

* This research was supported in part by the United States Air Force, through the Air Force Office of Scientific Research of the Air Research and Development Command, under Contract No. AF 49(638)- 213. Reproduction in whole or in part is permitted for any purpose of the United States Government.

$x_i + x_j$ in the cell (x_i , x_j), where addition is as defined earlier. A Latin square orthogonal to I, based on an orthogonal mapping α of G is derived by putting $\alpha x_i + x_j$ in the cell (x_i , x_j). The rows of this square are a permutation of the rows of I.

It was then shown that the search for orthogonal mappings was materially simplified by using a configuration based on the balanced incomplete block design with the parameters

$$v = b = 4t - 1, \quad r = k = 2t - 1, \quad \lambda = t - 1$$

or an equivalent Hadamard matrix of order $4t$. Using this method two sets of five mutually orthogonal Latin squares of order 12 were obtained. The mutually orthogonal Latin squares of a set, derived by using orthogonal mappings of a group enjoy the property D_0 [2] that all the members can be obtained from the basic square I by row-permutations alone. But following method of [1] we could generate for $v = 12$, sets of five mutually orthogonal Latin squares and no more. We wanted to investigate if it might be possible to increase the number in a set of mutually orthogonal Latin squares if we required the squares of a set to satisfy a property less restrictive than D_0 . This search was not successful. However, an interesting class of triplets of mutually orthogonal squares not possessing the property D_0 was discovered.

Let us regard each row of a square to consist of two *half-rows* — the first half of the row consists of the cells belonging to the $2t$ columns $x_1 , x_2 , \cdots , x_{2t}$, and the cells of the second half-row are those belonging to the columns $x_{2t+1} , \cdots , x_{4t}$. Instead of D_0 we now require our standard Latin square orthogonal to I to have its $4t$ first half-rows to be a permutation of the first half-rows of I and similarly its second half-rows a permutation of those of I.

2. *Use of double-maps of the group to generate orthogonal squares.*

Consider two (1, 1) mappings β and β^* (not necessarily orthogonal) of the elements of G, such that $\beta x_1 = x_1$ and $\beta^* x_1 = x_1$. Consider the square obtained by putting in the cell (x_i , x_j) the entry

$$\begin{aligned} &\beta x_i + x_j \quad \text{if} \quad 1 \le j \le 2t \\ &\beta^* x_i + x_j \quad \text{if} \quad 2t + 1 \le j \le 4t \end{aligned} \tag{2.1}$$

Theorem 2.1. A square whose cells are filled as in (2.1) is a Latin square if and only if

$$P_1(\beta x_i) = P_1(\beta^* x_i) \quad \text{for every } i.$$

Proof. Obviously, in every column all the elements of G occur. For a given x_i

$$\beta x_i + x_j \qquad 1 \le j \le 2t$$

is the set U of $2t$ elements of G such that

$$P_1(x) = P_1(\beta x_i) \quad \text{for every } x \text{ in } U,$$

and

$$\beta^* x_i + x_j \qquad 2t + 1 \le j \le 4t$$

is the set V of $2t$ elements x of G such that

$$P_1(x) = 1 + P_1(\beta^* x_i) = 1 + P_1(\beta x_i) \quad \text{for every } x \text{ in } V.$$

Hence, for a given x_i, all the elements of G will occur in row x_i.

Converse is easily proved.

Consider a Latin square which is based on a double-map (β, β^*) of G. Specifically, β and β^* satisfy

$$\beta x_1 = x_1, \qquad \beta^* x_1 = x_1 \quad \text{and} \quad P_1[\beta x_i] = P_1[\beta^* x_i] \tag{2.2}$$

for every x_i in G. We now require the condition that this Latin square $L(\beta, \beta^*)$ is orthogonal to I.

Let $S_0(\beta)$ denote the set of elements x of G such that $P_1(\beta x) = 0$ for every x in $S_0(\beta)$, i.e.,

$$S_0(\beta) = \{x : P_1(\beta x) = 0\}$$

Similarly, we define

$$S_1(\beta^*) = \{x : P_1(\beta^* x) = 1\}$$

and also

$$I_0(\beta) = \{\beta x : P_1(\beta x) = 0\} \qquad I_1(\beta^*) = \{\beta^* x : P_1(\beta^* x) = 1\}$$

It is easy to see that $S_0(\beta)$ and $S_1(\beta^*)$ are disjoint and $S_0(\beta) \cup S_1(\beta) = G$ and the same is true of $I_0(\beta)$ and $I_1(\beta^*)$. These are consequences of β and β^* satisfying (2.2).

On superimposing $L(\beta, \beta^*)$ on I we get $4t \times 4t$ ordered pairs of elements $\binom{x_i}{x_j}$ $1 \leq i, j \leq 4t$. With each pair is associated a difference $\theta = x_i - x_j$. If the two squares are mutually orthogonal, then it is known [3] that the $4t \times 4t$ pairs which are all different fall into $4t$ different classes—each class having $4t$ pairs. Any two members belonging to the same class have the same difference θ associated them and θ ranges over all the $4t$ elements of G. Given a pair $\binom{x_i}{x_j}$ which has the difference θ, the members belonging to its class are obtained as

$$\binom{x_i + x_u}{x_j + x_u} \qquad 1 \leq u \leq 4t$$

Consider the $4t \times 2t$ ordered pairs,

$$\text{(2.3) and} \quad \begin{aligned} &\binom{\beta x + x_1}{x + x_1}, \binom{\beta x + x_2}{x + x_2} \cdots \binom{\beta x + x_{2t}}{x + x_{2t}} \quad \text{for} \quad x \in S_0(\beta) \\ &\binom{\beta^* x + x_{2t+1}}{x + x_{2t+1}}, \cdots \binom{\beta^* x + x_{4t}}{x + x_{4t}} \quad \text{for} \quad x \in S_1(\beta^*) \end{aligned}$$

The $2t$ elements $(x_1, x_2, \cdots, x_{2t})$ of G, each occur as a top element in the $4t$ pairs (2.3). Now, if the $4t$ differences

$$\text{(2.4)} \qquad \begin{aligned} &x - \beta x \qquad x \in S_0(\beta) \\ &x - \beta^* x \qquad x \in S_1(\beta^*) \end{aligned}$$

generate all the $4t$ elements of G, then all the $4t \times 2t$ ordered pairs $\binom{x_u}{x_v}$ $1 \leq u \leq 2t$, $1 \leq v \leq 4t$ would be generated by (2.3). (2.4) is equivalent to saying that the mapping α defined by

$$\alpha: \begin{cases} x \to \beta x & x \in S_0(\beta) \\ x \to \beta^* x & x \in S_1(\beta^*) \end{cases} \tag{2.5}$$

is an orthogonal mapping of G. Similarly, we require that the mapping α^* defined by

$$\alpha^*: \begin{cases} x \to \beta x & x \in S_1(\beta) \\ x \to \beta^* x & x \in S_0(\beta^*) \end{cases} \tag{2.6}$$

is an orthogonal mapping of G, in order that the remaining $4t \times 2t$ ordered pairs $\binom{x_u}{x_v}$ $2t+1 \leq u \leq 4t$, $1 \leq v \leq 4t$ occur. Thus we have

Theorem 2.2. A Latin square $L(\beta, \beta^*)$ based on a double map (β, β^*) of G is orthogonal to I if β and β^* define two orthogonal mappings α and α^* as follows

$$\alpha: \begin{cases} x \to \beta x & x \in S_0(\beta) \\ x \to \beta^* x & x \in S_1(\beta^*) \end{cases}$$

$$\alpha^*: \begin{cases} x \to \beta x & x \in S_1(\beta) \\ x \to \beta^* x & x \in S_0(\beta^*) \end{cases}$$

It is easily seen that given two orthogonal mappings α and α^* of G such that

$$P_1(\alpha x) = P_1(\alpha^* x) \quad \text{for every } x \text{ in } G$$

one can always define a double-map (β, β^*) of G based on (α, α^*) in the manner of theorem 2.2, to generate a new Latin square $L(\beta, \beta^*)$ orthogonal to I and different from $L(\alpha)$ and $L(\alpha^*)$.

3. *Illustrations of double-maps*:

For $t = 3$, the mappings α and α^* given below are orthogonal mappings of $G(2, 6)$.

I :	00	01	02	03	04	05	10	11	12	13	14	15
α :	00	02	01	10	12	11	14	13	15	04	03	05
α^* :	00	02	04	12	11	13	14	10	15	03	05	01

Then β and β^* derived from α and α^* are as follows:

I :	00	01	02	03	04	05	10	11	12	13	14	15
β :	00	02	01	12	11	13	14	10	15	04	03	05
β^* :	00	02	04	10	12	11	14	13	15	03	05	01

The underlined elements of the maps α and α^* constitute β and the rest go to β^*. The Latin square based on the double-map β and β^* is given on the next page. It is easily verified that it is orthogonal to I.

$L(\beta, \beta^*)$ Orthogonal to I

β	00	01	02	03	04	05	10	11	12	13	14	15	β^*
02	02	03	04	05	00	01	12	13	14	15	10	11	02
01	01	02	03	04	05	00	14	15	10	11	12	13	04
12	12	13	14	15	10	11	00	01	02	03	04	05	10
11	11	12	13	14	15	10	02	03	04	05	00	01	12
13	13	14	15	10	11	12	01	02	03	04	05	00	11
14	14	15	10	11	12	13	04	05	00	01	02	03	14
10	10	11	12	13	14	15	03	04	05	00	01	02	13
15	15	10	11	12	13	14	05	00	01	02	03	04	15
04	04	05	00	01	02	03	13	14	15	10	11	12	03
03	03	04	05	00	01	02	15	10	11	12	13	14	05
05	05	00	01	02	03	04	11	12	13	14	15	10	01

FIGURE 1.

4. *Mutually orthogonal double-maps*

For $t = 3$, in [1] 10 sets of orthogonal mappings S_2 , $S_3 \cdots S_{11}$ were obtained. Number of mappings $n(S_i)$ in the sets S_i's are given below:

TABLE 1

i	2	3	4	5	6	7	8	9	10	11
$n(S_i)$	92	108	80	80	84	84	80	84	80	84

From the results in Section 2, it follows that by the method of double maps we can generate $n(S_i) \times n(S_i)$ Latin squares orthogonal to I using $n(S_i)$ mappings of set S_i .

Next the problem was one of matching for orthogonality 92×92 double maps (β_2 , β_2^*) from set S_2 , against 108×108 double-maps (β_3 , β_3^*) from set S_3 . The condition for (β_2 , β_2^*) to be orthogonal to (β_3 , β_3^*) is easily seen to be that the mappings γ and γ^*

$$\gamma: \quad \begin{aligned} &\beta_2 x \to \beta_3 x \quad \text{for } x \text{ satisfying } \quad P_1(\beta_3 x) = 0 \\ &\beta_2^* x \to \beta_3^* x \quad \text{for } x \text{ satisfying } \quad P_1(\beta_3^* x) = 1 \end{aligned}$$

and

$$\gamma^*: \quad \begin{aligned} &\beta_2 x \to \beta_3 x \quad \text{for } x \text{ satisfying } \quad P_1(\beta_3 x) = 1 \\ &\beta_2^* x \to \beta_3^* x \quad \text{for } x \text{ satisfying } \quad P_1(\beta_3^* x) = 0 \end{aligned}$$

are orthogonal mappings.

Using this condition for orthogonality, we obtained 228 pairs of mutually orthogonal double maps constructed one from each set of S_2 and S_3 . This

means including I we got 228 triplets of mutually orthogonal Latin squares of order 12, based on double maps one each from sets S_2 and S_3. However, in an extensive search by the computer, we could not get a set of four mutually orthogonal Latin squares based on double-maps which are not trivially orthogonal maps (α, α).

A triplet of mutually orthogonal double-maps of which one is I is given on the following page. α_2 and α_2^* are two orthogonal mappings from set S_2, and α_3 and α_3^* orthogonal mappings from S_3. Double maps (β_2, β_2^*) and (β_3, β_3^*) are derived respectively from (α_2, α_2^*) and (α_3, α_3^*) in the manner illustrated in Section 3.

I :	00	01	02	03	04	05	10	11	12	13	14	15
α_2 :	00	02	01	10	12	11	14	13	15	03	05	04
α_2^* :	00	04	01	12	10	13	11	15	14	03	05	02
α_3 :	00	11	04	15	02	13	01	14	05	12	03	10
α_3^* :	00	15	04	13	02	11	05	14	03	12	01	10
β_2 :	00	02	01	12	10	13	11	15	14	03	05	04
β_2^* :	00	04	01	10	12	11	14	13	15	03	05	02
β_3 :	00	15	04	13	02	11	01	14	05	12	03	10
β_3^* :	00	11	04	15	02	13	05	14	03	12	01	10

The mutually orthogonal Latin squares based on I, (β_2, β_2^*), and (β_3, β_3^*) are given in Figure 2. The symbols 00, 01, $\cdots$, 15 have been changed to 1, 2, $\cdots$, 12.

$L(I)$

1	2	3	4	5	6	7	8	9	10	11	12
2	3	4	5	6	1	8	9	10	11	12	7
3	4	5	6	1	2	9	10	11	12	7	8
4	5	6	1	2	3	10	11	12	7	8	9
5	6	1	2	3	4	11	12	7	8	9	10
6	1	2	3	4	5	12	7	8	9	10	11
7	8	9	10	11	12	1	2	3	4	5	6
8	9	10	11	12	7	2	3	4	5	6	1
9	10	11	12	7	8	3	4	5	6	1	2
10	11	12	7	8	9	4	5	6	1	2	3
11	12	7	8	9	10	5	6	1	2	3	4
12	7	8	9	10	11	6	1	2	3	4	5

$L(\beta_2, \beta_2^*)$

1	2	3	4	5	6	7	8	9	10	11	12
3	4	5	6	1	2	11	12	7	8	9	10
2	3	4	5	6	1	8	9	10	11	12	7
9	10	11	12	7	8	1	2	3	4	5	6
7	8	9	10	11	12	3	4	5	6	1	2
10	11	12	7	8	9	2	3	4	5	6	1
8	9	10	11	12	7	5	6	1	2	3	4
11	12	7	8	9	10	4	5	6	1	2	3
12	7	8	9	10	11	6	1	2	3	4	5
4	5	6	1	2	3	10	11	12	7	8	9
6	1	2	3	4	5	12	7	8	9	10	11
5	6	1	2	3	4	9	10	11	12	7	8

$L(\beta_2, \beta_2^*)$

1	2	3	4	5	6	7	8	9	10	11	12
12	7	8	9	10	11	2	3	4	5	6	1
5	6	1	2	3	4	11	12	7	8	9	10
10	11	12	7	8	9	6	1	2	3	4	5
3	4	5	6	1	2	9	10	11	12	7	8
8	9	10	11	12	7	4	5	6	1	2	3
2	3	4	5	6	1	12	7	8	9	10	11
11	12	7	8	9	10	5	6	1	2	3	4
6	1	2	3	4	5	10	11	12	7	8	9
9	10	11	12	7	8	3	4	5	6	1	2
4	5	6	1	2	3	8	9	10	11	12	7
7	8	9	10	11	12	1	2	3	4	5	6

FIGURE 2.

REFERENCES

1. R. C. BOSE, I. M. CHAKRAVARTI, AND D. E. KNUTH, "On methods of constructing sets of mutually orthogonal Latin squares using a computer—I". *Technometrics*, Vol. 2 (1960), pp. 507–516.
2. R. C. BOSE AND K. R. NAIR, "On complete sets of Latin Squares" *Sankhya*, Vol. 5 (1941), pp. 361–382.
3. R. C. BOSE, S. S. SHRIKHANDE, AND E. T. PARKER, "Further results on the construction of mutually orthogonal Latin squares and the falsity of Euler's conjecture". *Canadian J. Math.*, Vol 12 (1960), pp. 189–203.
4. H. B. MANN, "The construction of orthogonal Latin squares". *Ann. Math. Stat.*, Vol. 13 (1942), pp. 418–423.

Book Reviews

HAROLD F. DODGE and HARRY G. ROMIG, *Sampling Inspection Tables, Single and Double Sampling*, 2nd. ed. Revised and Expanded. New York: John Wiley and Sons, Inc., 1959. xi, 224 pp. quarto, $8.

The ground-breaking work of Dodge and Romig in the development of sampling inspection came to its culmination with publication of their sampling tables in the *Bell System Technical Journal* in January, 1941. These were later put together with mathematical and other explanatory material and published by Wiley in book form in 1944. The *Tables* have for many years been used widely in industry and their contributions to the spread of sound quality control methods was in good measure responsible for the authors being individually awarded the Shewhart Medal, Dodge in 1949, Romig in 1953.

The sampling plans of the Dodge-Romig *Tables* are minimum total inspection plans, i.e., they are so designed as to yield, for a specified process average, a minimum long range average amount of inspection, including the inspection of rejected lots as well as inspection of the sample. They are indexed with reference to the Average Outgoing Quality Limit (AOQL) and the Lot Tolerance Per cent Defective (LTPD) (the latter for a Consumer's Risk of 10%) and are classified with respect to lot size and process average. They thus stand in contrast to the *Mil. Std.* tables, subsequently published by the Department of Defense, which are indexed with respect to Average Quality Level (AQL) only.

The numerical tables of the revised edition are the same as the numerical tables of the original edition except for minor changes in lay-out. Since the new volume has approximately double the dimensions of the old (the new pages are now 8-1/2″ x 11″), the tables are set in larger size print. The term Process Average and the symbol % also now appear at the head of *each* column instead of only in the margin. Both these changes will increase the accuracy and ease of using the tables. A minor change that the reviewer did not like was the omission from the LTPD tables themselves of any statement of the Consumer's Risk.

The major changes in the volume as a whole are the addition of Operating Characteristic curves and the enlargment of the introductory chapter to include detailed directions of how to choose a sampling plan. The content of the original text in Chapters 1, 2, and 3 remains substantially unchanged.

One hundred and nine pages of the new edition are used to present the OC curves of the tabulated AOQL plans. Generally four charts are shown on each 8-1/2″ x 11″ page and the charts frequently contain 5 or 6 curves each. Like the numerical tables, the curves are indexed by AOQL's and grouped by lot sizes. OC curves for single sampling plans are based on exact binomial formulas, but those for double sampling plans are based on the Poisson approximation when sample sizes are large.

In the text, methods are given for deriving approximate OC curves for the tabulated LTPD plans from the curves given for the AOQL plans. Eight pages are also used to present a general set of OC curves for single sampling plans for $c = 0, 1, 2, 3$ and $n \leq 500$. These curves are indexed with respect to c and grouped according to values of n.

In their discussion of OC curves, Dodge and Romig distinguished Type A and Type B OC curves. Type A OC curves are defined as curves showing the probability of accepting a lot as a function of *lot* quality. Type B OC curves are defined as curves showing the probability of accepting a lot as a function of *process* quality. The authors point out that Type B curves are identical with Type A curves for $N = \infty$, since if lots can be viewed as random samples from a controlled process with fraction defective p, their lot qualitites will be binomially distributed and samples from such lots may be treated as if they were actually samples taken directly from an infinite universe having a fraction defective p.

The OC curves given in the *Tables* are all Type B OC curves, i.e., they show what percentage of lots would be accepted if the *process* is running (in control) at the p level. The

authors point out that "while not so designated in the original published papers of Chapters 1 and 2, the Consumer's Risk, as used there, is more specifically, the Consumers' A Risk, and the Producers Risk, as used there, is the Producers' B Risk," (p. 56.)

This publication of the OC curves for the Dodge-Romig AOQL plans is a welcome addition to the original tables and will make the plans more generally useful. Reiteration of their emphasis upon AOQL type plans may also possibly induce the federal government to reconsider the AOQL aspects of *Mil. Std.* plans.

As noted above, Dodge and Romig base the "optimum" character of their plans squarely on the attainment of a minimum quantity of total inspection. It was therefore a little disappointing to the reviewer that they did not at some place in the revised edition discuss recent approaches to the problem of sampling inspection that seek to minimize overall cost, including the cost of passing a defective unit as well as the cost of inspection. The authors, however, do give the user of the tables the following advice. "In choosing a value of LTPD (or AOQL), consider and compare the cost of inspection with the economic loss that would ensue if quality as bad as the LTPD were accepted often (or if the *average level* of quality were greater than the AOQL). Even though the evaluation of economic loss may be difficult, relative values for different levels of per cent defective may often be determined." (p. 6.)

Acheson J. Duncan

RONALD A. HOWARD "Dynamic Programming and Markov Processes," (Technology Press and Wiley: 1960) pp. 136, $5.75.

Statisticians, industrial engineers, operations research analysts and other technical personnel are frequently called upon either to provide counsel or to execute the solution of problems involving models or operational systems. This short, easy to read, and highly stimulating book will be helpful to such people in framing and solving certain types of mathematical models.

The author's purpose is to develop a computational technique—similar to dynamic programming—for solving particular problems arising from Markov processes. These problems pose the following question: how to maximize the attainment of some "reward", which is linked to a Markov process, through the manipulation of a parameter (the available policies). The iterative method suggested by Howard offers the advantage of being adequately descriptive, computationally feasible and usually providing the solution in a small number of iterations.

Chapter 1 deals with discrete-time Markov processes. The author presents a simple example of a Markov process in the behavior of a frog in a lily pond. Each lily pad is a state of the system. At a given point in time the frog occupies the i'th pad (the i'th state where $i = 1, 2, \cdots N$) and he leaps from pad to pad at random. If he moves at the end of equal discrete-time intervals and if his transition time is instantaneous, we have a discrete-time Markov process. The probability of moving from state i to state j for every state may be known. It is thus possible to define a transition matrix which consists of all transition probabilities.

At an early stage of the analysis the author introduces a generating function called a z-transform. The z-transform provides both limiting state probabilities (i.e. the probability distribution of state occupancy after a large number of leaps, or time intervals) and an expression for evaluating the transient behavior of the system. This is followed by a discussion of the application of the z-transform to systems portraying multichain and periodic behavior.

Chapter 2 introduces a reward variable, which is assumed to be a function of the transition made by the system. This implies that the reward depends upon the transition matrix. In the case involving our frog, the reward structure could emerge from a set of friendly wagers between another party and oneself and hence, from my point of view, some rewards could be positive, others negative. In the author's models, the dimension of the reward is arbitrary, so one can deal with dollars or physical units.

After introducing a definition of "expected reward", which is dependent upon the initial state, the z-transform is applied to the definition to facilitate deriving total "expected rewards" as a function of time. In this case the process is assumed to have but one recurrent chain.

Suppose now that we are able to influence the movements of the frog (by perhaps prodding him with an electric shock). If, while he is sitting on the i'th pad, the charge is applied to his head, his response may be described by a new vector of transition probabilities. There may be k alternative body locations to which the charge may be applied. For each pad which he may occupy we arbitrarily assign a particular surface of the frog's anatomy for prodding, i.e. to each state one, and only one, surface is assigned. This N-component vector of assignments is called a "policy". There exist, of course, a number of policies, depending upon N and the number of alternative body surfaces available for each state.

Given the expected reward function for the system the author illustrates how the total expected reward can be maximized by choosing an appropriate policy. In chapter 3 this is accomplished by iteration. But this is only feasible if n, the number of time intervals, is small. Suppose n is large. Then we turn to Chapter 4 where Howard's main contribution, the policy-iteration method, is clearly explained.

There is one other difference introduced at this point: for a large number of time intervals it is more useful to maximize *average* gain per transition rather than *total* gain. With this change in maximizing criterion, and assuming both a large n and that the limiting state probabilities for each policy are independent of the initial state, the policy-iteration method is derived. The optimal policy is a vector, with components specifying the decision to be made in each state, and it will usually be found in a small number of iterations.

Simple but clearly explained applications of the method to problems of taxicab operation, baseball strategy, and automobile replacement constitute the subject matter of Chapter 5. The last application is particularly revealing and useful since so many replacement problems in industry possess similar logical structures. The reader will find in this chapter a ready computational technique for solving such problems.

Chapter 6 deals with an extension of the policy-iteration method to Markov processes with more than one recurrent chain. In such cases the limiting state probabilities depend upon the starting state and the policy-iteration method of Chapter 4 is revised to accommodate this fact.

In Chapter 7 we find a revision of the method to reflect a different criterion: maximizing the present value of the expected reward. In other words, the rewards are discounted to give effect to the fact that immediate rewards are valued more highly than future rewards. Some insights are provided indicating numerical values of the discount factor that are important and under what conditions the no-discount solution of previous chapters will be equivalent to the discount solution.

Finally, the methods are extended to embrace continuous-time processes—i.e., the time between transitions is random. For the study of such Markov processes the Laplace transform is introduced to supplant the z-transform. A policy-iteration method, analogous to the one derived for discrete-time processes, is developed for solving long-duration continuous time decision problems involving rewards. Fortunately, the two procedures are computationally equivalent, so that a computer program developed for one procedure may be employed to solve problems posed by the other. The continuous-time-policy iteration method, like the discrete-time method, can be modified to handle discounting of expected rewards.

An excellent feature of this book is the liberal use of carefully chosen illustrative examples, literally one to support each new idea.

The book is designed for advanced courses in operations analysis or systems design courses at the graduate level in schools of engineering. But it will also serve as an excellent reference for technical personnel who are called upon to study Markov processes, mathematical models based upon Markov processes, or similarly structured problems.

William Beranek

Statistical Programs for High Speed Computers

Assembled for Technometrics by

FRED C. LEONE

Case Institute of Technology

Readers are requested to send brief descriptions of computer programs of interest to statisticians to Dr. Fred C. Leone, Director, Statistical Laboratory, Case Institute of Technology, 10900 Euclid Avenue, Cleveland 6, Ohio for consideration for publication in *Technometrics*. Each computer program announcement is to contain:

1. A brief problem description, 2. The type of computer for which the program is applicable, 3. The author's name and address, 4. A brief description of the program including: Limitations, auxiliary equipment requirements, statements on accuracy, availability of sample problems, running time estimates and programs storage requirements. All programs submitted for publication must be available for distribution subject only to nominal costs for the reproduction and mailing of program tapes, cards, etc. Inquiries about published programs should be addressed to the authors. Comments and critiques of the published programs may be sent to Dr. Leone and will be published when considered appropriate.

Multiple Regression by Stepwise Procedure, No. 6015 3112,—(1959)— IBM 650 With Alphabetic

This program performs a multiple linear regression by a stepwise procedure providing a number of intermediate equations by adding one variable at a time. The variable added is the one that makes the greatest improvement in "goodness of fit". Only significant variables are included in the final regression. It can be used with up to 33 independent variables. No limits on the number of sets of data. Data must be coded so that the largest sum of squares is less than 10^7. Weights can be assigned to each set of data. Procedure provides for entering and removing variables at specified F levels of significance. This program was successfully run and a sample problem and flow diagram are available. M. A. Efroymson and O. Schricker, Computing Center Esso Research and Engineering Co., Florham Park, P. O. Box 209, Madison, N. J.

A Modified Principle Axis Factor Analysis Solution, No. 0018, August 1958, IBM 650

An iterative solution, for symmetric matrices up to order 47 convergence can be obtained on the vectors of the right orthonormal (eigenvectors) and elements of the basic structure delta matrix (square root of the eigenvalues). The orthonormal vectors are also multiplied by the corresponding delta elements to produce vectors of the principal axes factor loading matrix. Successive residual matrices may be punched out, and factor extraction continued until the sum of the squared elements of the delta matrix equals or exceeds some specified value.

The program is not limited to positive definite matrices and has been designed to automatically process negative delta elements. When all systematic variance has been extracted, negative deltas may automatically follow, thereby reducing the sum of the elements so that the stop value is never reached. Under these conditions, extraction may continue until a machine overflow occurs when division by zero is attempted. In addition to its employment in the investigation of basic structure, this program may also be used for the rapid determination of the rank of a symmetric matrix. This program can be employed to factor symmetric matrices up to order 47. Only the infradiagonal and diagonal elements are loaded, and diagonal elements may or may not be unity. The program is self-restoring, and need not be reloaded to be reused. Information on wiring for input and output forms are provided. This program was successfully run. No diagram nor sample problem is available. Director, Div. of Counseling and Testing

Services, University of Washington, Seattle, Washington.

Statistical Multiple Correlation and Regression Analysis, UNIVAC I

This program can handle up to 90 variables, any number of these may be independent or dependent. The original input matrix could be unityped or prepared on the computer from some previously converted data. Card layout forms and a card-to-tape plugboard wiring diagram for the preparation of the input are given. A predicting equation of the form

$$Y^1 = A + B_1X_1 + B_2X_2 + \cdots + B_NX_N$$

will be found for each Y variable. The constant term A and the regression coefficients B appear as elements of output matrices, regardless of the total 16 output matrices, regardless of the total of dependent variables. The matrices contain such answers as the means and standard deviations of the original variables, the beta coefficients, the correlation coefficients between each of the independent variables and between the independent and dependent variables, the multiple correlation coefficients squared, the residuals, the standard errors of estimate, the "t" values, the regression coefficients, and the estimated value of Y. This routine is completely automatic, requiring no type-ins or operator intervention except in case of situations requiring reruns. This program was successfully run and a flow diagram is available. Director, Research and Technology Dept., National Tube Division, United States Steel Corp., Pittsburgh, Pa.

Data Screening No. 1, BIMED No. 015, March 10, 1960, IBM 709

This program is for testing input data which may be used for statistical analysis in later stages to find the form of their distributions, frequencies, and possible outliers. Input data may consist of n sets of data containing p variables with the following limitations:

(1) $n \leq 2000$; (2) $p \leq 50$;

and (3) $n \times p \leq 20{,}000$.

This program can perform a different transformation of any variables, if desired, according to the code specified in the Selection Card. A listing of the transformed and untransformed data is provided if desired. In order to process numbers of any size, the data for two selected variables are converted into the standardized form before they are screened and cross tabulated. Extreme values in the data are identified and printed in the output. This program was successfully run and a sample problem is available. Professor W. J. Dixon, Div. of Biostatistics, Dept. of Preventive Medicine and Public Health, School of Medicine, University of California at Los Angeles, Los Angeles 24, California.

Data Screening No. 2, BIMED No. 016, March 10, 1960, IBM 709

This program computes (1) frequency distributions, (2) means, variances, standard deviations and standard errors of means for a specified set of variables conditioned on one variable within the set, (3) correlation coefficients for the specified set of variables. The input data consists of n sets of p variables with the following restrictions on n and p: $n \leq 650$; $p \leq 30$.

This program can perform a different transformation of any variable, if desired, according to the code specified in the Transformation Card. This program has been successfully run and a sample problem is available. Professor W. J. Dixon, Div. of Biostatistics, Dept. of Preventive Medicine and Public Health, School of Medicine, University of California at Los Angeles, Los Angeles 24, California.

Least Squares Curve Fitting, 6.0.501, IBM 650

Given a set of independent data (x_i, y_i) $i = 1, 2 \cdots, N$, the program determines polynomials of degree less than a specified M. The weights, w_i may either be specified as part of the empirical data, or if an explicit function is available, may be calculated within the program. Fits up to the 24th degree if using all powers, 48th degree if using only even powers, 49th degree if using only odd powers for at most 100 experimental data points. All arithmetic is in floating point. The program does not need equally spaced values of the independent variable which is an important feature of the set of orthogonal polynomials used. The coefficients of orthogonal polynomials do not depend on the maximum degree specified. This program was successfully run and a sample problem and flow diagram are available.

Director, Computing Center, University of Rochester, Rochester, New York.

Least Square Curve Fitting on Unequally Spaced Points with Orthogonal Polynomials, No. 1006, April 1958, IBM 650, SOAP II Complier (Developed at Case Inst. of Tech.)

Given a set of empirical data (x_u , y_u) $u = 1, 2, \cdots , m$, the program determines polynomials of degree less than or equal to 10. The program does not need equally spaced values of the independent variable which is an important feature of the set of orthogonal polynomials used. The coefficients of orthogonal polynomials do not depend on the maximum degree specified. This program was successfully run and a flow diagram is available. Director, Computing Center, Case Institude of Technology, Cleveland 6, Ohio.

Fitting on Unequally Spaced Points with Orthogonal Polynomials, No. 1032, Sept. 1960, BU 220 ALGOL (Burroughs Algebraic Compiler)

This program will find the coefficients of the polynomial

$$y = c_0 + c_1x + c_2x^2 + \cdots + c_nx^n$$

which gives the best fit in a least squares sense to data points (up to 500 pairs) where the maximum value of n does not exceed fifteen. The data points may be arbitrarily spaced. The algorithm does not use the "normal" equations approach but employs orthogonal polynomials. This program was successfully run and a sample problem and flow diagram are available. Director, Computing Center, Case Institute of Technology, Cleveland 6, Ohio.

Analysis of Variance, IBM 650 Basic

The programs developed perform an Analysis of Variance on any of five types of experiment designs: one way classification of variates, randomized block, latin squares, split plot and split-split plot. An attempt has been made to allow for flexibility of input and output by making provisions that allow the input data to be coded in either fixed or floating-point form. The following are limitations:

(1) The maximum number of treatments in a one way classification of variates is 100.

(2) The maximum number of replicates or treatments allowable in a randomized block experiment is 100.

(3) The maximum latin square allowable is 100 × 100.

(4) The restrictions on the split plot program are:
   (a) The number of replicates $\leq$100.
   (b) The number of whole plot treatments $\leq$100.
   (c) The number of sub-plot treatments $\leq$100.
   (d) The product of whole plot treatments times sub-plot treatments $\leq$300.

(5) The restrictions on the split-split plot program are:
   (a) The number of replicates $\leq$32.
   (b) The number of whole plot treatments $\leq$32.
   (c) The number of sub-plot treatment $\leq$50.
   (d) The number of sub-sub-plot treatments $\leq$32.
   (e) (Whole plot treatments) × (sub-plot treatments) × (sub-sub-plot treatments) $\leq$200.

This program was successfully run and a sample problem and flow diagram are available. Miss Betty L. Takvan, Assistant Librarian, Numerical Analysis Laboratory, University of Arizona, Tucson, Arizona.

Analysis of Variance, BIMED No. 012, IBM 709, March 1960

This program is written for a complete factorial design. As many as 14 variables may be processed. There is no limitation on the number of replicates. The number of categories or levels of any variable cannot exceed 999, and input data in one replicate cannot exceed 20,000, namely:

$$L_1 \times L_2 \times \cdots \times L_n \leq 20{,}000$$

where $L_1 , L_2 \cdots , L_n$ are numbers of levels of the N variables, and $n \leq 14$. The program can also perform a transformation of the input data according to the codes specified in the control card. This program was successfully run and a sample problem is available. Professor W. J. Dixon, Div. of Biostatistics, Dept. of Preventive Medicine and Public Health, School of Medicine, University of California at Los Angeles, Los Angeles 24, California.

One and Two Way Analysis of Variance, UNIVAC I

This program is for computing the statistical parameters necessary for the Analysis of Variance of one and two factor experiments. The ANOVA table including F ratios, is printed out at the SCP (Supervisory Control Printer). In addition, column and row sums are put on an output tape.

The restrictions to the routine are:

1. $A + B \leq 300$
2. $N < 10^{11-d}$

where A is the number of levels for A factor, B is the number of levels for B factor, Y is the number of replications per cell, N is the total number of observations ($=A \times B \times Y$), d is the maximum number of digits in any observation.

A modified program is also available, for input containing 2 way experiments when it is desired to eliminate the printing of ANOVA tables at the SCP. This program is marked HSP (High Speed Printer) and follows the regular program on the instruction tape. For the 2 way experiments only, it will prepare the ANOVA table for the high speed printer. This program was successfully run. No sample problem nor diagram is available. Anthony Brescia, Program Library Services, Remington Rand Univac, Div. of Sperry Rand Corporation, 315 Park Ave., South, New York 10, N. Y.

Probit III, IBM 650, Indexing Accumulators and Floating Decimal Arithmetic.

This program fits a regression line to dosage-mortality or other quantal dose-response data. The method used is that of maximum likelihood described by Finney. The procedure is iterative and normally stops when two successive regression coefficients agree within 1/100th of their standard error. Different tolerances can be substituted. Maximum number of doses is 20. Maximum number of iterations 8. The program employs approximations for the normal curve which are significant to minimally 4 decimal places. Heterogeneity is tested by the χ^2 test. Normal operation of the program requires a 5% probability for significance tests. If heterogeneity is present, the heterogeneity factor is automatically included at the appropriate places in the computation. When the data are heterogeneous to such a degree that the regression is not significant, the machine will punch out a negative lower fiducial limit of the regression coefficient (when it is positive). The program was compiled by means of FORTRANSIT II.

This program was successfully run. No flow diagram nor sample problem is available. Director, Computation Center, University of Kansas, Lawrence, Kansas.

Statistical Extensions for Runcible I, Sept. 1959, IBM 650, Extensions are written in 650 machine language but are used as part of Runcible I programming system. (May be used in connection with any other Runcible Extensions.)

These statistical extensions, when used within a program written in Runcible I language, will compute the following statistics and/or probability functions:

Data information, mean, covariance, alpha 3, alpha 4, simple correlation coefficient, standard deviation, highest value, lowest value, variance, simple linear regression equation, cross product, correlation matrix, sum of variable, sum of squares, sum of third powers, sum of fourth powers, floating point random numbers, gamma function, cumulative normal probability, inverse cumulative normal probability, Poisson probability, cumulative Poisson probability, Binomial probability, cumulative Binomial probability, Binomial coefficients, Chi-square test, t test I & II, F test, simultaneous linear equation solver, matrix inverse, fixed point random numbers.

This report should be regarded as a supplement to the "RUNCIBLE I" and "RUNCIBLE EXTENSIONS—PART I" reports (available from Case Inst. of Tech.). This program was successfully run and a sample problem is available. Director, Statistical Laboratory, Case Institute of Technology, Cleveland 6, Ohio.

NOTICES

DoD HANDBOOK H-109

A new handbook in the Quality Control and Reliability series has been published by the Office of the Assistant Secretary of Defense (S&L). DoD Handbook H-109, "Statistical Procedures for Determining Validity of Suppliers' Attributes Inspection," may be purchased from the Superintendent of Documents, U. S. Government Printing Office, Washington 25, D. C. (price: 30 cents). This handbook provides appropriate statistical tests and tables of critical values for use in determining the validity of suppliers inspection records when sampling inspection by attributes is specified in the contractual document.

PROCEEDINGS, NATIONAL SYMPOSIA ON RELIABILITY AND QUALITY CONTROL

There have been six National Symposia on Reliability and Quality Control, co-sponsored by the American Society for Quality Control, the Institute of Radio Engineers, Inc., the American Institute of Electrical Engineers and the Electronic Industries Association. For each of the Symposia a paper-bound copy of PROCEEDINGS was printed. Within the past year stocks of four of these—First, Second, Third and Fifth—were depleted. These four have now been reprinted and the *PROCEEDINGS* of all six Symposia are now in stock for prompt shipment. Orders for desired volumes, at $5.00 net each volume postpaid, should be sent to the Editorial Department, Institute of Radio Engineers, 1 East 79th Street, New York 21, New York.

NEW YORK UNIVERSITY MEETING ON "MATHEMATICS AND STATISTICS FOR RELIABILITY"

The Electronics Division ASQC and the Section on Engineering and Physical Sciences, ASA, are sponsoring a meeting on "Mathematics and Statistics for Reliability Problems" to be held March 27 and 28, 1961 in New York City at New York University's Washington Square campus. There will be, in addition to a program of invited papers, several sessions for short contributed papers. These papers will be restricted to fifteen minutes presentation time. Persons having papers they feel merit presentation should send abstracts not to exceed 200 words to:

DR. WILLIAM A. GLENN, *Program Chairman*
Research Triangle Institute
Post Office Box 490
Durham, North Carolina

University of Nebraska Degrees in Statistics

The Mathematics Department of the University of Nebraska announces the introduction of a curriculum leading to the M.A. or M.S. and Ph.D. degrees in Statistics. A bachelor's degree program in Statistics is also under consideration at present. Courses are offered in Statistical Methods, Statistics for Engineers, Theory of Probability, Information Theory, Methods of Experimental Design, Stochastic Processes, Theory of Games and Statistical Decision Theory, and Topics in Probability and Statistics. Additional courses will be added as needed. A number of assistantships and fellowships are available to qualified students. For information concerning degree requirements and for information concerning assistantships, direct inquiries to Professor Bernard Harris, Department of Mathematics, University of Nebraska, Lincoln 8, Nebraska.

Summer Institute for College Teachers of Statistics

The National Science Foundation will sponsor a Summer Institute for College Teachers of Statistics at Iowa State University for the 11-week period from June 5 through August 18, 1961. The Departments of Statistics of three other universities, Kansas State, Utah State and the University of Wyoming, are cooperating with Iowa State's statistical center in presenting this institute.

Financial support in the form of stipends, dependency allowances and travel allowances will be awarded to 50 eligible applicants. All American college and university teachers who are, or who during the 1961-62 academic year will be, required to teach one or more courses in statistics as part of their regular assignments are eligible for consideration.

The institute is planned to provide additional basic training in statistics for present and prospective teachers who, though well-grounded in other fields, have limited backgrounds in statistics. Also it will provide more advanced courses and seminars designed to keep college and university teachers abreast of new developments.

Courses are scheduled in Statistical Methods, Theory of Statistics, Experimental Design, Survey Designs, Topics in Foundations of Probability and Statistics, and Intermediate Applied Decision Theory. In addition, an opportunity will be provided for those interested to observe a demonstration class in Principles of Statistics at the undergraduate level. The faculty will include the institute director, Dr. T. A. Bancroft, director of the Iowa State University Statistical Laboratory and head, Department of Statistics; Dr. R. J. Buehler, associate professor of statistics, Iowa State University; Dr. H. T. David, associate professor of statistics, Iowa State University; Dr. H. C. Fryer, head of the Department of Statistics and Statistical Laboratory director, Kansas State University; Dr. H. O. Hartley, professor of statistics, Iowa State University; the institute associate director, Dr. D. V. Huntsberger, associate professor of statistics, Iowa State University; and Dr. R. L. Hurst, head of the Department of Applied Statistics and Statistical Laboratory director, Utah State University. Guest lecturers will present a series of special seminars.

Requests for information or application forms should be addressed to: The

Director, Summer Institute in Statistics, 102 Service Building, Iowa State University, Ames, Iowa.

Southern Regional Graduate Summer Session in Statistics at Virginia Polytechnic Institute

The 1961 session of the Southern Regional Graduate Summer Session in Statistics will be held at the Virginia Polytechnic Institute, Blacksburg, Virginia, from June 15 to July 22, 1961. Blacksburg is 2100 feet high on the plateau in the Allegheny Mountains and offers a delightfully cool summer climate.

The Virginia Polytechnic Institute, Oklahoma State University, North Carolina State College, and the University of Florida have agreed to operate a continuing program of graduate summer sessions in statistics to be held at each institution in rotation. The program was instituted at Virginia Polytechnic Institute in the summer of 1954.

The 1961 session, like previous sessions under this program, is intended to serve: 1) teachers of statistics and mathematics; 2) professional workers who want formal training in modern statistics; 3) research and engineering personnel who want intensive instruction in basic statistical concepts and modern statistical methodology; 4) Public Health statisticians who wish to keep informed about advanced specialized theory and methods; 5) prospective candidates for graduate degrees in statistics; and 6) graduate students in other fields who desire supporting work in statistics.

The session will last six weeks and courses will carry five quarter hours of credit. Not more than two courses may be taken for credit at any one session. The summer work in statistics may be applied as residence credit at any of the cooperating institutions, as well as certain other universities, in partial fulfillment of the requirements for a graduate degree. The program may be entered at any session, and consecutive courses will follow in successive summers so that it would be possible for a student to complete the course work in statistics for a Master's degree in three summers. Students must satisfy the remaining requirements for course work and thesis at the institution where they are to be admitted to candidacy. The advanced courses may be accepted as part of the Ph.D. program of the participating institutions.

A limited number of fellowships will be available for applicants from certain specialized areas. Doctoral courtesy will be honored for those holding Ph.D. or M.D. degrees.

The courses to be offered in statistics in 1961 at the Virginia Polytechnic Institute are as follows: Statistical Methods, Sampling Theory; Applied Statistics for Engineers and Physical Sciences; Theory I, Probability; Theory II, Statistical Inference; Theory III, Theory of Linear Hypotheses; Non-parametric Methods; and Multivariate Methods.

A number of courses in advanced mathematics will be available during the Summer Session. For a complete listing please consult the University timetable. A series of Colloquia involving recent developments in statistical theory and methods will be conducted during the special Summer Session.

Requests for application blanks for the summer session and for fellowships should be addressed to:

DR. BOYD HARSHBARGER, *Head*
Department of Statistics
Virginia Polytechnic Institute
Blacksburg, Virginia.

PURDUE UNIVERSITY TEN DAY SHORT COURSE

Lafayette, Ind.—A ten-day short course on The Mathematical Techniques of Operations Research will be offered at Purdue University for the second time this summer, according to Prof. Paul H. Randolph, of the department of mathematics and statistics, director of the short course. The dates of the course are June 5-15, 1961.

The course has been designed specifically for statisticians, quality control analysts, engineers, and other technical personnel in industrial and management positions, Professor Randolph stated. Emphasis will be placed on the mathematical techniques of operations research and the application of these methods to current industrial and military problems.

These methods involve the construction of mathematical models representing the operation of industrial management or a military organization, and suggest the best solutions to problems involved. Among the topics to be discussed during the course are inventory control models, waiting line models, linear programming, simplex method, transportation methods, production scheduling models, search theory, cost-effectiveness studies and systems analysis.

Two outstanding leaders in this field have been selected as instructors for the short course. Dr. Albert Madansky, of the University of California at Los Angeles, did basic work in stochastic linear programming for the Rand Corporation, at Santa Monica, Calif. Dr. Bernard Lindgren, of the Institute of Technology at the University of Minnesota, and formerly a National Science Foundation scholar in Operations Research at Stanford University, is the co-author of a well known book on probability in statistics.

The course is sponsored by Purdue's Statistical Laboratory and Division of Adult Education. Further information may be obtained by writing to the Division of Adult Education.

Errata

"Some New Three Level Designs for the Study of Quantitative Variables"

G. E. P. BOX AND D. W. BEHNKEN, TECHNOMETRICS, Vol 2, No 4, Nov. 1960

This paper should carry the acknowledgement:

"This work was supported in part by the Department of the Army, Project No. 5B 99-01-004, Ordnance R & D Project No PB2-0001, OOR Project No. 1715, Contract No. DA-36-034-Ord 2297 through the Statistical Techniques Research Group, Princeton University."

"The Effect of Sequential Batching for Acceptance-Rejection Sampling Upon Sample Assurance of Total Product Quality"

M. HALPERIN AND G. L. BURROWS, TECHNOMETRICS, Vol 2, No 1, Feb. 1960

Formula 2.1, pg 21, should read:

$$Q=\frac{\left[\sum_1^K r_i+\frac{t_\alpha^2}{2}(1-f)\right]+t_\alpha\sqrt{(1-f)\left[(1-f)\frac{t_\alpha^2}{4}+\sum_1^K r_i-\left(\sum_1^K r_i\right)^2\Big/f\sum_1^K N_i\right]}}{\left[f+t_\alpha^2(1-f)\Big/\sum_1^K N_i\right]}$$

TABLES OF THE HYPERGEOMETRIC PROBABILITY DISTRIBUTION

Gerald J. Lieberman and Donald B. Owen. For the first time, the hypergeometric probability distribution is extensively tabulated and the results of exact machine computation are available. Of great use in mathematics, statistics, the social and physical sciences, and engineering. *Stanford Studies in Mathematics and Statistics, III.* $15.00

STUDIES IN ITEM ANALYSIS AND PREDICTION

Edited by Herbert Solomon. This integrated series of mathematical studies presents many new theoretical developments in both test design and the classification of individuals on the basis of responses to tests. *Stanford Mathematical Studies in the Social Sciences, VI.* About $8.75

MARKOV LEARNING MODELS FOR MULTIPERSON INTERACTIONS

Patrick Suppes and Richard C. Atkinson. A major attempt to bridge the gap between statistical learning theory and social psychology by applying stimulus-sampling techniques to social interaction situations. *Stanford Mathematical Studies in the Social Sciences, V.* $8.25

Order from your bookstore, please

Stanford University Press

BIOMETRICS

Journal of the Biometric Society

Vol. 17, No. 1 CONTENTS March 1961

Marginal Percentages in Multiway Tables of Quantal Data With Disproportionate Frequencies — *F. Yates*

Some Classification Problems with Multivariate Qualitative Data — *William G. Cochran and Carl E. Hopkins*

The Standardization of Tuberculin-Hypersensitivity — *M. Stone and R. A. Bruce*

Optimum Estimation of Gradient Direction in Steepest Ascent Experiments — *Samuel H. Brooks and M. Ray Mickey*

A Stochastic Study of the Life Table and Its Applications III. The Follow-Up Study with the Consideration of Competing Risks — *Chin Long Chiang*

The Spearman Estimator for Serial Dilution Assays — *Eugene A. Johnson and Byron Wm. Brown, Jr.*

The Fitting of a Generalization of the Logistic Curve — *J. A. Nelder*

Combined Analysis of Balanced Incomplete Block Designs with Some Common Treatments — *M. V. Pavate*

Generalized Asymptotic Regression and Non-Linear Path Analysis — *Malcolm E. Turner, Robert J. Monroe, and Henry L. Lucas*

On the Analysis of Split-Plot Experiments — *H. Leon Harter*

Queries and Notes

Confidence Intervals for Recombination Experiments with Micro-organisms — *A. W. Kimball*

Estimation of Proportion from Zero-Truncated Binomial Data — *G. N. Wilkinson*

Book Reviews

E. J. Gumbel: Statistics of Extremes — *B. F. Kimball*

A. Vessereau: Methodes Statistiques en Biologie et en Agronomie — *P. Robinson*

A. G. Love, E. L. Hamilton and I. L. Hellman; Tabulating Equipment and Army Medical Statistics — *W. W. Holland*

Biometrics is published quarterly. Its objects are to describe and exemplify the use of mathematical and statistical methods in biological and related sciences, in a form assimilable by experimenters. The annual non-member subscription rate is $7. Inquiries, orders for back issues and non-member subscriptions should be addressed to:

BIOMETRICS
Department of Statistics
The Florida State University
Tallahassee, Florida

Journal of the
AMERICAN STATISTICAL ASSOCIATION

VOLUME 55 NUMBER 292

DECEMBER 1960

TABLE OF CONTENTS

Indeterminism in Science and New Demands on Statisticians
Jerzy Neyman

On Finite Sample Distributions of GCL Identifiability Test Statistics ..... *Robert L. Basmann*

A Note on the Limiting Relative Efficiency of the Wald Sequential Probability Ratio Test ..... *Robert E. Bechhofer*

Bivariate Exponential Distributions ..... *E. J. Gumbel*

On the Exact Variance of Products ..... *Leo A. Goodman*

Circular Error Probabilities ..... *H. Leon Harter*

On Conditional Expectations of Location Statistics ... *Robert V. Hogg*

Effect of Bias on Estimates of the Circular Probable Error
P. B. Moranda

Market Growth, Company Diversification and Product Concentration 1947-1954 ..... *Ralph L. Nelson*

A New Binomial Approximation for Use in Sampling from Finite Populations ..... *P. J. Sandiford*

Internal Migration Statistics for the United States
Everett S. Lee and Anne S. Lee

Bibliography on Simulation, Gaming, Artificial Intelligence and Allied Topics ..... *Martin Shubik*

For further information, please contact:

American Statistical Association
1757 K Street, N. W.
Washington 6, D. C.

THE ANNALS OF MATHEMATICAL STATISTICS

Vol. 31, No. 4—December, 1960

Contents

Our Silver Anniversary ..... *Allen T. Craig*
Simplex-Sum Designs: A Class of Second Order Rotatable Designs Derivable from those of First Order ..... *G. E. P. Box and D. W. Behnken*
Third Order Rotatable Designs in Three Dimensions ..... *Norman R. Draper*
A Third Order Rotatable Design in Four Dimensions ..... *Norman R. Draper*
Some Aspects of Weighing Designs ..... *Damaraju Raghavarao*
Random Allocation Designs I: On General Classes of Estimation Methods ..... *A. P. Dempster*
A Mixed Model for the Complete Three-Way Layout with Two Random-Effects Factors ..... *J. P. Imhof*
Tests for Regression Coefficients when Errors are Correlated ..... *M. M. Siddiqui*
Mixed Model Variance Analysis with Normal Error and Possibly Non-Normal Other Random Effects, Part I: The Univariate Case ..... *S. N. Roy and Whitfield Cobb*
Mixed Model Variance Analysis with Normal Error and Possibly Non-Normal Other Random Effects, Part II: The Multivariate Case ..... *S. N. Roy and Whitfield Cobb*
An Associated Polynomial for Least Square Approximations ..... *Richard Warren*
Asymptotic Rate of Discrimination for Markov Processes ..... *Lambert H. Koopmans*
Equalitiies for Stationary Processes Similar to an Equality of Wald ..... *Shu-Teh Chen Moy*
Multivariate Chebyshev Inequalities ..... *Albert W. Marshall and Ingram Olkin*
On Deviations of the Sample Mean ..... *R. R. Bahadur and R. Ranga Rao*
On the Independence of a Sample Central Moment and the Sample Mean ..... *R. G. Laha, E. Lukacs and M. Newman*
Order Statistics of Partial Sums ..... *J. G. Wendel*
Probability Distributions Related to Random Mappings ..... *Bernard Harris*
Some Asymptotic Results for a Coverage Problem ..... *Max Halperin*
Statistical Programming ..... *D. F. Votaw, Jr.*
The Life Distribution and Reliability of a System with Spare Components ..... *Donald F. Morrison and H. A. David*
Maximizing the Probability that Adjacent Order Statistics of Samples from Several Populations form Overlapping Intervals ..... *Richard Cohn, Frederick Mosteller, John W. Pratt, and Maurice Tatsuoka*
Normalizing the Noncentral t and F Distributions ..... *Nico F. Laubscher*
Probability Content of Regions Under Spherical Normal Distributions, II: The Distribution of the Mean in Normal Samples ..... *Harold Ruben*
Tables of Range and Studentized Range ..... *H. Leon Harter*
Asymptotic Formulae for the Distribution of Hotelling's Generalized T_0^2 Statistics, II ..... *Koichi Ito*
On Unbiased Estimation ..... *L. Schmetterer*
Two-Stage Experiments for Estimating a Common Mean ..... *Donald Richter*
Rank-Sum Tests for Dispersions ..... *A. R. Ansari and R. A. Bradley*
A Relationship between Hodges' Bivariate Sign Test and a Non-Parametric Test of Daniels' ..... *Bruce M. Hill*
Minimax Sequential Tests of Some Composite Hypotheses ..... *M. H. DeGroot*
Inverse Binomial Sampling Plans When an Exponential Distribution is Sampled with Censoring ..... *Jack Nadler*

Notes:

A Conservative Property of Binomial Tests ..... *H. A. David*
An Optimum Property of Regular Maximum Likelihood Esitmation ..... *V. P. Godambe*

The purpose of the Institute of Mathematical Statistics is to encourage the development, dissemination and application of mathematical statistics. Membership dues, which include a subscription to the *Annals of Mathematical Statistics* are $10.00 per year for residents of the United States or Canada and $5.00 a year for residents of other countries. Inquiries regarding membership to the Institute should be sent to the Secretary: G. E. Nicholson Jr., Department of Statistics, University of North Carolina, Chapel Hill, North Carolina.

BIOMETRIKA

Volume 47, Parts 3 and 4 DECEMBER 1960

CONTENTS

Memoirs:

LESLIE, P. H. AND GOWER, J. C. The properties of a stochastic model for the predator-prey type of interaction between two species.

BLOM, GUNNAR. Hierarchical birth and death processes I. Theory.

BLOM, GUNNAR. Hierarchical birth and death processes II. Applications.

GLENN, W. A. A comparison of the effectiveness of tournaments.

PEARCE, S. C. Supplemented balance.

FOLKS, JOHN LEROY AND KEMPTHORNE, OSCAR. The efficiency of blocking in incomplete block designs.

HAIGHT, FRANK A. Queueing with balking, II.

DAVID, H. A. AND PEREZ, CARMEN A. On comparing different tests of the same hypothesis.

FARLIE, D. J. G. The performance of some correlation coefficients for a general bivariate distribution.

MCFADDEN, J. A. Two expansions for the quadrivariate normal integral.

LAHA, R. G. AND LUKACS, E. On a problem connected with quadratic regression.

MCCULLOUGH, ROGER S., GURLAND, JOHN AND ROSENBERG, LLOYD. Small sample behaviour of certain tests of the hypothesis under variance heterogeneity.

SUKHATME, BALKRISHNA V. Power of some two-sample non-parametric tests.

EWAN, W. D. AND KEMP, K. W. Sampling inspection of continuous processes with no autocorrelation between successive results.

BLYTH, COLIN R. AND HUTCHINSON, DAVID W. Table of Neyman-shortest unbiased confidence intervals for the binomial parameter.

BENNETT, B. M. AND HSU, P. On the power function of the exact tests for the 2×2 contingency table.

SIMPSON, J. A. AND WELCH, B. L. Table of the bounds of the probability integral when the first four moments are given.

SEVERO, NORMAN C. AND ZELAN, MARVIN. Normal approximation to the chi-square and non-central F probability functions.

BARTON, D. E., DAVID, F. N. AND O'NEILL, ANNE F. Some properties of the distributions of the logarithm of non-central F.

LINDLEY, D. V., EAST, D. A. AND HAMILTON, P. A. Tables for making inferences about the variance of a normal distribution.

BARTON, D. E., DAVID, F. N. AND MERRINGTON, M. Tables for the solution of the exponential equation, $\exp(-a) + ka = 1$.

Miscellanea:

Contributions by G. E. Bardwell, B. R. Bhat and J. Gani, D. R. Cox, R. N. Curnow, N. Gilbert, W. M. Harper and J. A. Macdonald, N. L. Johnson and D. H. Young, M. G. Kendall, A. N. Kshirsagar, H. Linhart, T. A. Ramsubban.

Reviews: *Other Books Received:*

The subscription, *payable in advance*, is 54/—(or $8.00) per volume (including postage). Cheques should be made payable to *Biometrika*, crossed "*a/c Biometrika Trust*" and sent to The Secretary, Biometrika Office, University College, London, W.C.1. All foreign cheques must be drawn on a Bank having a London agency.

Issued by THE BIOMETRIKA OFFICE, University College, London

INDUSTRIAL QUALITY CONTROL

• VOL. XVII, No. 5, NOVEMBER 1960

Conducting In-Plant Training Sessions in Statistics
J. L. Jaech and W. F. Stevenson

Training for Statisticians—Industry's Need.....*Truman L. Koehler*

The Role of Quality Control as Moderator of Applied Science In the Corporation.....*Leslie E. Simon*

Acceptance Sampling Demonstration with a Bead Box
Frank Cullen and Stanley Kalin

A Graphical Method for Determining the Effect of the Isolated Factor on the Total Span in the Span Plan Analysis
James R. Gilman

• VOL. XVII, No. 6, DECEMBER 1960

The Legal Framework for Inspection Under Government Contracts
Albert Green

Quality Control from Front Door to Back Door.....*Sigmund P. Zobel*

A Mixed Model Factorial Experiment in Testing Electrical Connectors
F. Lemus

The Experiment, the Design, the Analysis.....*Charles R. Hicks*

A Six Step Quality Control Program.....*Harley Wehrwein*

Sampling Plans for Determining Compliance of Electronic Parts Reliability Requirements.....*Irving B. Altman*

• VOL. XVII, No. 7, JANUARY 1961

S Q C vs. Intuitive Inspection.....*Chester T. Raymo*

Critique of Vendor Quality Rating Systems.....*Leroy Folks*

Quality Control in the Construction Industry.....*Joseph J. Waddell*

A Decentralized Team Approach to Quality and Reliability
T. E. Smith and A. W. Wortham

Industrial Quality Control is published monthly by the American Society for Quality Control, Inc. All correspondence concerning membership in the society and subscriptions to the journal should be addressed to American Society for Quality Control, Inc., Rm 6185 Plankinton Bldg., 161 West Wisconsin Ave., Milwaukee, 3, Wisconsin.

INTERNATIONAL
JOURNAL OF ABSTRACTS
STATISTICAL THEORY AND METHOD

COVERAGE OF THE JOURNAL

The aim of this journal is to give complete coverage of published papers in the field of statistical theory (including associated aspects of probability and other mathematical methods) and new published contributions to statistical method.

There are approximately two hundred and fifty journals published in various parts of the world which are wholly or partly devoted to the field of statistical theory and method and which are brought within the scope of this journal of abstracts.

In addition to the vast array of journal literature, abstracts of the special collections of papers as published in reports of conferences symposia and seminars, together with the published reports of experiment and other research stations, are aalso included.

FORMAT OF THE JOURNAL

The abstracts are about 400 words long—the recommendation of UNESCO for the "long" abstract service: they are in the English language although the language of the original paper is indicated on the abstract together with the name of the abstractor. In addition, the address of the author is given in sufficient detail to facilitate conttact in order to obtain further details or request an offprint.

A scheme of classification has been developed for the abstracts that is flexible and facilitates the transfer of code numbers to punched cards. Each abstract has two classification numbers: the primary number in heavy type to indicate the basic topic of the paper and the secondary number in brackets to take account of the most important cross-reference. A unique aspect of this journal is that the pages are colour-tinted according to the main sections of the classification. This method of colour-coding the pages provides a distinctive and powerful visual aid to the identification of abstracts in whatever manner the journal is filed for reference.

THIS JOURNAL is prepared by an international editorial organisation under a General Editor in association with a Managing Editor for the American and Pacific Area. The General Editor works in conjunction with the Research Techniques Unit (London School of Economics) and the Managing Editor is on the staff of the Institute of Statistics (North Carolina State College).

GENERAL EDITOR

Dr. Wm. R. Buckland

c/o 55 Broadway, London, S.W. 1, England

MANAGING EDITOR (AMERICA AND PACIFIC)

Prof. R. L. Anderson

North Carolina State College, Raleigh, N.C., U.S.A.

Annual subscription, £5 (U.S.A. & CANADA $16.00)
Single number, 30s (U.S.A. & CANADA $4.50)

PUBLISHED QUARTERLY FOR
THE INTERNATIONAL STATISTICAL INSTITUTE BY
OLIVER AND BOYD LTD.

Journal of the

ROYAL STATISTICAL SOCIETY

Series B (Methodological) Volume 22, No. 2, 1960

CONTENTS

Models in the analysis of variance (with Discussion) . . . . . R. L. PLACKETT

Discrete stochastic processes in population genetics (with Discussion) W. F. BODMER

A problem of delayed service—I . . . . . A. BEN-ISRAEL & P. NAOR

A problem of delayed service—II . . . . . A. BEN-ISRAEL & P. NAOR

On the transient behaviour of a simple queue . . . . . P. D. FINCH

Queueing at a single serving point with group arrival . . . . . B. W. CONOLLY

On some extensions of Bayesian inference proposed by Mr. Lindley. . R. A. FISHER

A method of estimation of missing values in multivariate data suitable for use with an electronic computer . . . . . S. F. BUCK

Estimation of parameters of a multivariate normal population from truncated and censored samples . . . . . NAUNIHAL SINGH

Necessary restrictions for distributions *a posteriori* . . . . . D. A. SPROTT

Weight of evidence, corroboration, explanatory power, information and the utility of experiments . . . . . I. J. GOOD

The logistic process: tables of the stochastic epidemic curve and applications EDWIN MANSFIELD & CARLTON HENSLEY

Some notes on pistimetric inference . . . . . A. D. ROY

Evaluation of determinants, characteristic equations and their roots for a class of patterned matrices . . . S. N. ROY, B. G. GREENBERG & A. E. SARHAN

Bounds for the expected sample size in a sequential probability ratio test M. N. GHOSH

A note on tests of homogeneity applied after sequential sampling . . . D. R. COX

The interaction algorithm and practical Fourier analysis: an addendum . I. J. GOOD

Approximate solutions of Green's type for univariate stochastic processes H. E. DANIELS

The Wilcoxon tests and non-null hypotheses . . . . . G. B. WETHERILL

Index to Volume 22 (1960)

ROYAL STATISTICAL SOCIETY, 21 BENTINCK ST., LONDON W. 1

Back issues of the Society's Journal, Series A (General) and Series B (Methodological), are available. Prices and other particulars of the Journal and of other publications still available may be obtained from the Secretary of the Society.

PREPARATION OF MANUSCRIPTS

Manuscripts should be submitted to the office of the editor: J. S. Hunter, Mathematics Research Center, U. S. Army; The University of Wisconsin; Madison, Wisconsin. Each manuscript should be typewritten, double spaced, with wide margins at sides, top, and bottom. The original should be submitted with two additional copies, on paper that will take corrections. Dittoed or mimeographed papers are acceptable only if completely legible. Footnotes should be avoided and replaced by remarks in the text, or placed in an appendix. Preferably, references in the manuscript should appear as (Jones, A. B., 1958). and again later in alphabetical order in a list of references. Alternatively, references may be numbered, e.g. [1], as they appear in the manuscript and be listed in this sequence in the list of references. In the reference list, each reference should contain, in the order indicated, the name and initials of the author followed by those of the co-authors, date of publication, title of reference, source, volume number and page. References to books should include publisher's name and location.

Figures, charts, and diagrams should be professionally drawn on plain white paper or tracing cloth in black India ink twice the size they are to be printed. A full page diagram, in print, measures 7.25 × 4.75 inches.

As far as possible, formulas should be typewritten and symbols not available on a typewriter carefully inserted in ink. Authors are asked to keep in mind the typographical difficulties of complicated mathematical formulae. The difference between capital and lower-case letters should be clearly shown; care should be taken to avoid confusion between such pairs as zero and the letter O, the numeral 1 and the letter l, numeral 1 used as superscript and prime ($'$), alpha and a, kappa and k, mu and u, nu and v, eta and n, etc. Subscripts or superscripts should be clearly below or above the line. Bars above groups of letters (e.g., $\overline{\log x}$) and underlined letters (e.g., $\underline{x}$) are difficult to print and should be avoided. Symbols are automatically italicized by the printer and should not be underlined on manuscripts. Boldface letters may be indicated by underlining with a wavy line on the manuscript; boldface subscripts and superscripts are not available. Complicated exponentials should be represented with the symbol exp particularly when appearing in the text, that is,

$$\exp\,[(a^2 + b^2)^{1/2}] \text{ should be used in place of } e^{(a^2+b^2)^{1/2}}.$$

In writing square roots the fractional exponent is preferable to the radical sign. Fractions in the body of the text (and when possible in displayed expressions) and fractions occurring in the numerators or denominators of fractions are preferably written with the solidus; thus

$$(a + b)/(c + d) \text{ rather than } \frac{a + b}{c + d}.$$

Authors will ordinarily receive only galley proofs. Fifty offprints without covers will be furnished free. Costs for additional reprints and covers can be furnished on request.

TECHNOMETRICS

The purpose of **Technometrics** is to contribute to the development and use of statistical methods in the physical, chemical and engineering sciences. This objective places a high premium on succinct communication among the physical scientist, engineer, statistician and mathematician. The journal will accept for publication papers describing new statistical techniques expected to be useful in these sciences, papers illustrating the application of known statistical methods to new or novel environments, expository or tutorial papers on particular statistical methods, and papers dealing with the philosophy and problems of applying statistical methods to research, development, design and performance. Brief descriptions of problems requiring solution and short technical notes will also be accepted for publication. Letters to the Editor, signed by the author and limited in length will be published when they are considered timely and appropriate. All papers should contain a short but clear summary of contents and conclusions, an expository section containing numerical examples whenever practicable, and appropriate additional sections relating to technical derivations.

Subscription Rates

The annual subscription rate for members of the American Society for Quality Control or the American Statistical Association is $6.00 a year. The annual subscription rate for non-members is $8.00 a year.

Members of the sponsoring societies, the American Society for Quality Control and the American Statistical Association, may subscribe to the journal while paying their annual dues or by check or money order made out to Technometrics and mailed to:

TECHNOMETRICS
Post Office Box 587
Benjamin Franklin Station
Washington 4, D. C.

All non-member subscriptions should be mailed to this address. Communications concerning changes of address, subscriptions, back numbers, etc., should also be sent to this address. Whenever possible, a copy of the address taken from an issue of the periodical should accompany a change of address request. All subscription fees are payable in the currency of the United States of America.

Communication concerning *membership* in either of the sponsoring societies should be sent to that society.

Second class postage paid at Richmond, Va.